## REMARKS

Claims 1, 2, 4, 6, 7, 19, 20, 21, 22, 23, and 26 are currently pending and under examination. Claims 3, 5, 8-18, 24, 25, and 27-41 have been withdrawn. Claims 1, 4, 6, 7, 19, 20, and 26 are currently amended. Applicants request entry of the claim amendments presented here. Applicants request that the remarks presented below be considered, and the amended claims be found in condition for allowance.

### *Claim Objections*

Claim 1 is objected to for the spelling of "produceable." Claim 1 has been amended to recite "producible."

Claims 4, 6, 7, 19-23, and 26 are objected to under 37 CFR §1.75(c) for allegedly improper form. Claims 4, 6, 7, 19, 20, and 26 have been amended to change their dependency.

Claims 19-23 are objected to for depending on non-elected claims. Claims 19 and 20 have been amended to depend on elected claims.

### *Claim Rejections Under 35 USC §112*

Claims 2, 4, 6, 7, 19, 20, 21, 22, 23, and 26 are rejected under 35 USC §112, second paragraph, as allegedly indefinite. Applicants respectfully submit that the claim amendments presented herewith have overcome these rejections.

Claim 2 is rejected for allegedly lacking antecedent basis for the limitation "the functional domains" in line 1. Claim 4 is rejected as dependent on rejected Claim 2. Applications point out that Claim 1 has been amended to recite "functional domains," thereby rendering these rejections moot.

Claim 6 is rejected as allegedly indefinite with respect to the limitation "said domains" in line 2. Applicants point out that Claim 6 has been amended to recite "said functional domains" in line 2, thereby obviating this objection.

Claim 7 is rejected as allegedly indefinite with respect to the limitation "said domains" in line 2. Applicants point out that Claim 7 has been amended to recite "said functional domains" in line 2, thereby obviating this objection.

Claim 19 is rejected as allegedly lacking antecedent bases for the limitation "said constant domain" in lines 1-2. Applicants respectfully traverse. Claim 19 is dependent on Claim 1, which recites "the constant CL-domain of an immunoglobulin light chain." Thus, Claim 1 provides sufficient antecedent basis for the limitation "said constant domain of an immunoglobulin light chain" recited in Claim 19.

Claim 20 is rejected as allegedly lacking sufficient antecedent basis for the limitation "said constant immunoglobulin domains" in lines 1-2 and "said functional receptor-ligand domains" in lines 2-3. Claim 20 has been amended to recite "said immunoglobulin constant domains". Applicants point out that Claim 1, as currently amended, recites "said immunoglobulin constant domains" at lines 9-10. Thus, Claim 1 provides sufficient antecedent basis for "said constant immunoglobulin domains" in line 2 of Claim 20. Claim 20 has further been amended to recite "functional domains having receptor or ligand function." Claim 1, as currently amended, recites ""functional domains having different receptor or ligand functions" at lines 7-8. Thus, Claim 1 provides sufficient antecedent basis for the limitation "functional domains having receptor or ligand function " recited in amended Claim 20.

Claims 21-23 are rejected because they depend on rejected based claim 20. Applications respectfully submit that amended Claim 20 is in condition for allowance and therefore, this rejection of dependent Claims 21-23 is improper and should be withdrawn.

Claim 26 is rejected as allegedly indefinite with respect to the limitation "CH1 domain is limited to a histidine tag" in line 2. Applicants have amended Claim 26 as suggested by the Examiner, such that Claim 26 now recites "CH1 domain is linked to a histidine tag."

Applicants submit that all rejections under 35 USC §112, second paragraph, have been overcome and these rejections should be withdrawn.

***Claim Rejections Under 35 USC §102(b)***

Claims 1, 2, 4, 6, 7, 19, 20, 21, 22, and 26 are rejected under 35 USC §102(b) as allegedly anticipated by Müller *et al.* (1998, *FEBS Lett.* 422:239-264). Applicants traverse this rejection for the reasons presented below.

To anticipate a claim, the reference must teach every element of the claim. "'A claim is anticipated only if each and every element as set forth in the claim is found, either expressly or inherently described, in a single prior art reference.'" MPEP § 2131, quoting *Verdegaal Bros. v. Union Oil Co. of California*, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987). Applicants submit that Müller *et al.* does not disclose a multifunctional compound producible in a mammalian host cell, nor does Müller *et al.* disclose at least two different functional domains that lack an intrinsic affinity for one another. Therefore, the rejection of claims 1, 2, 4, 6, 7, 19 to 22 and 26 of the present application under 35 §102(b) over Müller *et al.* is improper and should be withdrawn.

***The Müller et al. reference does not disclose a multifunctional compound producible in a mammalian host cell.***

Amended Claim 1 recites "a multifunctional compound, producible in a mammalian host cell as a secretable and fully functional heterodimer of two polypeptide chains . . ." In contrast, Müller *et al.* discloses bispecific miniantibodies that functionally assemble in *E. coli*.

According to the Office Action, the recitation of "produceable (producible) in a mammalian host cell as a secretable and fully functional heterodimer of two polypeptide chains" in Claim 1 allegedly does not appear to limit either the function and/or structure of the claimed multifunctional compound. (Office Action, page 6 line 21 to page 7, line 1). It was further noted that, "the recitation of a process limitation in claim 1 is not viewed as positively limiting the claimed product absent a showing that the process of making recited in claim 1 imparts a novel or unexpected property to the claimed product, as it is assumed that equivalent products are obtainable by different routes." (Office Action, page 7, lines 14-20).

To assert equivalence, an Examiner "must provide rationale or evidence tending to show inherency." MPEP §2122. Applicants point out that no rationale or evidence has been proposed to support the argument that the bispecific miniantibodies of Müller *et al.* are equivalent to the multifunctional compound of Claim 1. This argument is improper and should be withdrawn.

Nonetheless, Applicants comply with the alleged burden to establish a patentable distinction between the claimed product and the product of the Müller *et al.* reference by showing that the process of making the multifunctional compound of Claim 1 imparts at least one "novel or unexpected property" to the claimed compound.

> The term "fully functional" means, in accordance with the present invention, that the compounds of the invention secreted by mammalian host cells into the culture supernatant in contrast to e.g. proteins expressed as inclusion bodies in E.coli do not require any protein refolding after purification; all subunits of the compounds of the invention are correctly folded and thus express their specific functions simply by being expressed in and secreted by mammalian host cells. (Specification as filed, page 5, lines 17-22)

Because it has been shown that the process of making the multifunctional compound of Claim 1 imparts at least one "novel or unexpected property" to the claimed compound, Applicants submit that the *E coli*-assembled bispecific antibodies of Müller *et al.* are not equivalent to the claimed multifunctional compounds producible in a mammalian host cell.

***The Müller et al. reference does not disclose multifunctional compounds comprising different functional domains that lack intrinsic affinity for one another.***

Claim 1 of the present application, as amended, recites a multifunctional compound comprising functional domains having different receptor or ligand functions, "wherein further at least two of said different functional domains lack an intrinsic affinity for one another." In contrast, Müller *et al.* discloses bispecific antibodies having two different scFv-fragments. In the specification, it was noted that the results of Müller *et al.* using Cl and CH1 to support dimerization of scFv fragments is not surprising because "scFv-fragments are also known to form dimers with each other even without the support of any special dimerization domains." (Specification as filed, page 7, lines 24-26). That is, scFv-fragments are a class of molecules well documented in the literature as forming dimers due to mutual affinity, even without the support of a dimerization domain. Accordingly, the bispecific antibodies disclosed by Müller *et al.* are not the claimed multifunctional compounds comprising functional domains having different receptor or ligand functions, "wherein further at least two of said different functional domains lack an intrinsic affinity for one another."

Applicants present the following support for the argument that the bispecific antibodies containing scFv fragments disclosed in Müller *et al.* are *not* the claimed multifunctional compounds. Kortt *et al.*, cited in the specification (page 7, lines 28-29) and submitted here as Exhibit A, states:

> "The purified, monomeric scFv was found to form dimers and higher-molecular-mass aggregates at a protein concentration greater than 5 mg/ml. Recently, the presence of dimers and higher-molecular-mass multimers was reported for several scFv molecules [citation to Griffiths *et al.*, 1993, *EMBO J.* 12:725-734] and the utilization of scFv dimers as bispecific molecules was demonstrated [citation to Holliger *et al.*, 1993, *PNAS USA* 90:6444-6448]." (Exhibit A: Kortt *et al.*,1994, *Eur. J. Biochem.* 221, page 151 right column, line 19 to page 152, left column, line 3.)

Griffiths *et al.*, cited in the specification at page 7, line 29, and attached as Exhibit B, describes the natural dimerization properties of scFv-fragments in the following section:

> "***The scFv fragments can form both monomers and dimers in solution***
> Soluble antibody fragments were purified from bacterial supernatants by affinity chromatography, by binding of C-terminal peptide tag to the mAb 9E10 (Munro and Pelham, 1986; Clackson *et al.*, 1991; Marks *et al.*, 1991). After ultrafiltration, the fragments were further

> purified by FPLC gel filtration (Pharmacia), and detected on-line both by UV absorption (280 nm) and by binding to antigen immobilized on a sensor chip in BIAcore (Pharmacia Biosensor AB) (Jönsson *et al.*, 1991; Jönsson and Malmqvist, 1992). This showed that the scFv fragments emerged in two peaks, corresponding in size to monomers and dimers (Figure 4). . . . The dimers run as monomers on non-reducing SDS gels (Laemmli, 1970) (not shown) and are therefore not linked by disulphide bonds. As two peaks are seen in gel filtration, it appears that in this case the monomers and dimers do not interconvert rapidly (for discussion and references for gel filtration equilibria see Jones *et al.*, 1985). Presumably the dimers are scFv fragments interlocked through the flexible linker joining the heavy and light chains, or with the heavy chain of one scFv molecule associated with the light chain of the other." (Exhibit B: Griffiths *et al.*,1993, *EMBO J.* 12, page 728, right column, line 1 to 25, emphasis added).

Applicants wish to point out that scFv-dimers as shown in Figure 4 of Griffiths *et al.* (Exhibit B) were found by gel filtration in every example they studied, lending support to the conclusion that dimer formation is a general property of scFv-fragments. Therefore, Kortt *et al.* (Exhibit A) and Griffiths *et al.* (Exhibit B) support the argument that scFv-fragments such as those found in the bispecific antibodies of Müller *et al.*, have an "intrinsic affinity for one another". As disclosed in Neiba *et al.* cited at page 7, line 33, and attached as Exhibit C, the structural determinant for this dimerization has been identified as hydrophobic amino acids at the C-termini of the variable domains that are buried in intact IgG or Fab molecules but become exposed in single chain scFv constructs. (Nieba *et al.* 1997, *Protein Eng.* 10: 435-444).

***In conclusion***, Applicants submit that Müller *et al.* does not teach every element of Claim 1 because it has been shown that the bispecific antibodies disclosed by Müller *et al.* are not the equivalent of the claimed multifunctional compound of Claim 1, and because it has been shown that Müller *et al.* does not disclose a multifunctional compound comprising functional domains having different receptor or ligand functions, "wherein further at least two of said different functional domains lack an intrinsic affinity for one another" as recited in Claim 1. Applicants conclude that Müller *et al.* does not anticipate Claim 1, or Claims 2, 4, 6, 7, 19, 20, 21, 22, and 26 which incorporate the claim limitations of Claim 1. Therefore, the rejection of Claims 1, 2, 4, 6, 7, 19, 20, 21, 22, and 26 under 35 USC §102(b) over Müller *et al.* is improper and should be withdrawn.

***Claim Rejections Under 35 USC §103, Obviousness***

Claims 1, 2, 4, 6, 7, 19, 20, 21, 22, 23, and 26 are rejected under 35 USC §103, as allegedly unpatentable over Müller *et al.* (1998, *FEBS Lett.* 422:239-264) in view of Plückthun and Pack (1997, *Immunotechnology* 3:83-105). It is argued that the Müller *et al.* reference allegedly discloses a multifunctional compound with all the structural limitations of Claims 1, 2, 3, 6, 7, 19, 20, 21, 22, and 26, for reasons given for the rejection under 35 USC §102(b) discussed above, while the Plückthun and Pack reference allegedly teaches the use of hinge regions, in particular the upper hinge from human Ig3 [*sic*]. (Office Action, page 9, lines 9-22). Using these disclosures, it is argued that it would have been obvious to one of ordinary skill in the art to substitute the linkers of Müller *et al.* with the upper hinge region of human IgG3 taught by Plückthun and Pack, to make a multifunctional compound as claimed. Applicants submit that the Examiner has not established a *prima facie* case of obviousness by combining these references and therefore, the rejection under 35 U.S.C. § 103(a) is improper and should be withdrawn.

To establish a *prima facie* case of obviousness, three basic criteria must be met. First, there must be some suggestion or motivation, either in the references themselves or in the knowledge generally available to one of ordinary skill in the art, to modify the reference or to combine reference teachings. Second, there must be a reasonable expectation of success. Finally, the prior art reference, or references when combined, must teach or suggest all the claim limitations. *See*, MPEP §§ 2142, 2143. In the present case, there is no suggestion or motivation to modify or combine the references teachings in the manner proposed, there is no reasonable expectation of success from the proposed combination, and the proposed combination does not teach all the claim limitations. Because basic criteria for a *prima facie* case of obviousness have not been met, rejection of Claims 1, 2, 4, 6, 7, 19, 20, 21, 22, 23, and 26 under 35 U.S.C. §103 is improper and should be withdrawn.

***The cited references provide no suggestion or motivation to combine references to make the claimed invention, and no reasonable expectation of success.***

Obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. *In re Fine*, 837 F.2d 1071, 5 USPQ 2d 1596 (Fed.

Cir. 1988); *In re Jones* 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). See generally, MPEP §2143, especially §2143.01.

Applicants traverse the argument put forth in the Office Action that Müller *et al.* provides motivation to use a human sequence (Office Action, page 9, lines 13-15), and Plückthun and Pack provide and motivation to use a human IgG hinge region (see, Office Action, page 10, lines 1-3). Although Müller *et al.* discloses bispecific antibodies expressed in *E. coli* using what appear to be human sequences, Müller *et al.* does *not* teach or suggest that their bispecific antibodies are producible in a mammalian host cell, as claimed herein. Absent any teaching that their *E. coli*-expressed bispecific antibodies would be producible in a mammalian host cell, there is neither a suggestion or motivation to combine the cited references to include an IgG3 hinge, nor is there any reasonable expectation of success that such a combination would produce the claimed invention.

In addition, Applicants point out that there is no suggestion or motivation to combine the cited references to reject 1, 2, 4, 6, 7, 19, 20, 21, or 26, as the Plückthun and Pack reference would only be applicable to those claims that are directed to use of IgG hinge regions (*e.g.*, Claims 22 and 23).

***The cited references, alone or in combination, do not teach all the claim limitations***

To establish *prima facie* obviousness of a claimed invention, all the claim limitations must be taught or suggested by the prior art MPEP §2143.03, citing *In re Royka*, 490 F.2d 981, 180 USPQ 580 (CCPA 1974). Here, the cited references do not teach or suggest the claimed multifunctional compound comprising functional domains having different receptor or ligand functions, "wherein further at least two of said different functional domains lack an intrinsic affinity for one another." Müller *et al.* discloses bispecific antibodies containing two scFv-fragments which, as discussed above, are known to have an intrinsic affinity for one another. Therefore, neither Müller *et al.*, nor the combination of Müller *et al.* and Plückthun and Pack, teaches all the claim limitations of the claimed multifunctional compound.

***In conclusion,*** because the cited references, alone or in combination, fail to meet the criteria for a *prima facie* case of obviousness, Applicants submit that the claimed invention is not obvious in light of the cited references. Therefore, rejection of Claims 1, 2, 4, 6, 7, 19, 20, 21, 22, 23, and 26 under 35 U.S.C. §103 is improper and should be withdrawn.

## CONCLUSION

Claims 1, 2, 4, 6, 7, 19, 20, 21, 22, 23, and 26 are currently pending and under examination. Claims 3, 5, 8-18, 24, 25, and 27-41 have been withdrawn. Claims 1, 4, 6, 7, 19, 20, and 26 are currently amended. Applicants request entry of the claim amendments presented here. Applicants request that all outstanding rejections be withdrawn in light of the remarks presented herein, and the amended claims be found in condition for allowance.

Applicants believe that no fees are due. If any fees are due, please charge any additional fees to Deposit Account No. 50-2212.

If the Examiner has any questions about this communication, she is encouraged to contact the Applicants' attorney at (858) 509-4093.

Respectfully submitted,

Date: Oct 13, 2004

Donna O. Perdue, Ph.D.
Reg. No. 51,166

Pillsbury Winthrop LLP
11682 El Camino Real, Suite 200
San Diego, CA 92130-2593
Telephone: (858) 509-4093
Facsimile: (858) 509-4010

Eur. J. Biochem. 221, 151–157 (1994)




# Recombinant anti-sialidase single-chain variable fragment antibody

## Characterization, formation of dimer and higher-molecular-mass multimers and the solution of the crystal structure of the single-chain variable fragment/sialidase complex

Alexander A. KORTT[1], Robyn L. MALBY[2], J. Bruce CALDWELL[1], L. Clem GRUEN[1], Neva IVANCIC[1], Michael C. LAWRENCE[2], Geoffrey J. HOWLETT[3], Robert G. WEBSTER[4], Peter J. HUDSON[1] and Peter M. COLMAN[2]

[1] Commonwealth Scientific and Industrial Research Organisation, Division of Biomolecular Engineering, Parkville, Australia
[2] Biomolecular Research Institute, Parkville, Australia
[3] Department of Biochemistry, University of Melbourne, Parkville, Australia
[4] Department of Virology, St Jude Children's Research Hospital, Memphis, USA

(Received September 10/October 21, 1993) – EJB 93 1388/3

The single-chain antibody variable fragment (scFv), with a 15-residue polypeptide linker $(Gly_4Ser)_3$, of monoclonal antibody NC10 was expressed in *Escherichia coli* and purified to homogeneity. This scFv molecule, refolded from 6 M guanidine hydrochloride, was predominantly a monomer of 27 kDa and was stable on storage at 4° and 20°C. At higher protein concentrations (≈ 5 mg/ml) dimer and higher-molecular-mass multimers were formed and freezing enhanced this aggregation. The dimer was not stable and dissociated to monomer at 20°C with a half-life of approximately 8 days. The higher-molecular-mass multimers and dimer dissociated to monomer in 60% ethylene glycol. Both the monomer and dimer were active and with tern N9 sialidase yielded complexes of 276 kDa and 569 kDa, respectively, indicating that four scFv molecules bound/sialidase tetramer and that the dimer was bivalent and cross-linked two sialidase tetramers. Binding studies at low concentrations and using radiolabelled scFv indicated that the binding affinity of the dimer was approximately twofold higher than that of the monomer, and the binding affinities of the scFv were similar to that of the parent NC10 antigen-binding fragment (Fab) molecule. A complex between tern N9 sialidase and NC10 scFv was crystallized and the structure of the complex was solved at 0.3-nm resolution by X-ray diffraction. Comparison of this scFv/sialidase structure with the parent Fab/sialidase structure revealed that the modes of attachment of scFv and Fab to sialidase were very similar. There was no discernible electron density for the peptide linker joining the variable heavy ($V_H$) and variable light ($V_L$) chains. A close interaction between two symmetry-related scFv suggests that they may have crystallized as dimers.

Monoclonal antibodies with their unique specificity and affinity are used widely as immunodiagnostic and therapeutic reagents [1]. The cloning of antibody genes [2] and recent advances in design, construction and expression of antibody domains in mammalian [3, 4] and bacterial cells [5, 6] have led to the production of a new range of recombinant antibody fragments for diagnostic and therapeutic applications. Recombinant antigen-binding (Fab) and variable (Fv) fragments have been expressed in bacterial systems but present the problem of correct assembly of the separate heavy and light chains in the extracellular environment [7, 8].

Single-chain antibody variable fragments (scFv), in which the variable heavy ($V_H$) and variable light ($V_L$) domains are covalently joined with a polypeptide linker, fold in the correct conformation and have improved stability compared to Fv fragments [9, 10]. The single-chain Fv is the smallest fragment to be produced that shows equivalent binding affinity to the parent Fab fragment [11–14]. The amino acid composition and the length of the linker have been investigated and a flexible 15-residue peptide linker of the type $(Gly_4Ser)_3$ has been used in the construction of several scFv molecules [15, 16]. The ability to engineer stable scFv fragments, including the attachment of molecular ligands as immunotoxins and the generation of bivalent molecules [17, 18] will enhance their diagnostic and therapeutic applications.

A scFv fragment of antibody NC10, which recognizes an epitope on influenza virus sialidase, was recently cloned and expressed in *Escherichia coli* [14]. The $V_H$ and $V_L$ domains were linked with the 15-residue peptide described above and also contained a hydrophilic octapeptide (FLAG) [19] at the C-terminus of the $V_L$ domain as an affinity label to aid in detection of the scFv. The purified, monomeric scFv was found to form dimers and higher-molecular-mass aggregates at a protein concentration greater than 5 mg/ml. Recently, the presence of dimers and higher-molecular-mass multimers

*Correspondence to* A. A. Kortt, CSIRO, Division of Biomolecular Engineering, 343 Royal Parade, Parkville, Victoria 3052, Australia
*Fax:* +61 3 347 5481.

*Abbreviations.* scFv, single-chain antibody variable fragment; Fab, antigen-binding fragment of antibody; $V_H$ and $V_L$, variable domains of antibody heavy chain and light chains; PelB, pectate lyase secretion-signal peptide from *Erwinia cartovora*; rms, root mean square.

was reported for several scFv molecules [20] and the utilization of scFv dimers as bispecific molecules was demonstrated [21].

The complex between this NC10 scFv and tern N9 sialidase has been crystallized for three-dimensional structural studies [14]. The structures of the complexes between N9 sialidase, isolated from tern [22] and whale [23], and the Fab fragments of monoclonal antibodies NC41 [24, 25] and NC10 [26], which recognize overlapping epitopes on sialidase, have been determined. Sialidase is a homotetramer with circular fourfold symmetry and both of these complexes show four Fab molecules bound to the tetrameric antigen [24, 26]. The protein sequences of the tern and whale N9 sialidases differ at 14 residues [27] but these substitutions have no effect on the Fab/sialidase interface [25] although they may influence the apparent binding constant for the interaction with NC41 Fab [28].

Here we report the purification of NC10 scFv, with a 15-residue linker, the properties of the monomer-dimer interaction and the solution of the crystal structure of the tern N9 sialidase/scFv complex which reveals a possible interaction between the $V_H$ domains to form a dimer.

## MATERIALS AND METHODS

The design of NC10 scFv from the $V_H$ and $V_L$ genes of NC10 (a monoclonal antibody with specificity toward N9 sialidase), the construction of the scFv gene and its expression in *E. coli* using the secretion vector pPOW [29] was recently described [14]. SDS/PAGE analysis of bacterial cell fractions showed that the NC10 scFv was present in the periplasm as insoluble protein aggregates and associated with the bacterial membrane fraction [14].

The NC10 scFv was solubilized by extraction of the bacterial membrane fraction with 6 M guanidine hydrochloride/0.1 M Tris/HCl, pH 8.0 by stirring at 4°C for 16 h. Insoluble membrane debris was removed by centrifugation and the guanidine hydrochloride removed by dialysis against 0.025 M Tris/HCl, 0.1 M NaCl, pH 8.0, at 4°C. Protein which precipitated during dialysis was removed by centrifugation and the soluble fraction, which contained about 50% of the extracted scFv, was concentrated approximately eightfold by ultrafiltration (Amicon stirred cell, YM10 membrane). The concentrate was applied to a Sephadex G-100 column (60 cm×2.5 cm) equilibrated with 0.025 M Tris/HCl, 0.1 M NaCl, pH 8.0, and fractions which contained protein were analysed on SDS/PAGE. The scFv was located by staining with Coomassie brilliant blue G or by Western blot analysis using anti-FLAG M2 antibody (IBI). SDS/PAGE was performed in 0.75-mm or 1.5-mm slabs [12% (mass/vol.) polyacrylamide] in a Bio-Rad Mini-Protean II gel apparatus. For Western blots, proteins were transferred to nitrocellulose (Schleicher and Schuell) with 0.025 M Tris/glycine, pH 8.5, containing 20% methanol as the transfer buffer. The detection reagent was sheep anti-mouse-horseradish-peroxidase conjugate as the second antibody and 4-chloro-1-naphthol as the colour reagent. Fractions containing scFv were pooled, concentrated and dialysed against 0.025 M Tris/HCl, pH 8.0, and chromatographed on a Mono-Q column (Pharmacia-LKB). The purified scFv was finally concentrated to about 5 mg/ml and stored frozen.

Monomeric and multimeric species of the purified scFv were separated by size-exclusion HPLC on a Superose 6 or 12 (Pharmacia), or Superdex 75 (Pharmacia) column in 0.025 M Tris/HCl, 0.1 M NaCl, pH 8.0. The columns were calibrated with the following standard proteins: thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), bovine serum albumin dimer (132 kDa), bovine serum albumin (66 kDa), ovalbumin (45 kDa) and soybean trypsin inhibitor (21.5 kDa). The stability of purified monomer and dimer species was monitored by size-exclusion HPLC on the Superdex 75 column.

Tern and whale N9 sialidases were purified as described [30] and concentrated to 5 mg/ml for crystallization studies. The binding activity of the scFv monomer and dimer was assessed qualitatively by size-exclusion HPLC of complexes formed with tern N9 sialidase. The affinity constants for the complexes of tern and whale N9 sialidase with NC10 scFv monomer and dimer were determined by sedimentation equilibrium using a Beckman Airfuge as described [31]. Trace amounts ($10^{-9}$ M) of $^{125}$I-labelled scFv monomer and dimer were used with sialidase concentrations of $10^{-6}$–$10^{-9}$ M for these binding experiments. The molecular mass of the complexes of tern N9 sialidase with NC10 scFv monomer and dimer was determined in 0.05 M sodium phosphate, 0.15 M NaCl, pH 7.4 by sedimentation equilibrium in a Beckman model XLA ultracentrifuge. Partial specific volumes of 0.702 ml/g and 0.704 ml/g were calculated for sialidase and scFv from the amino acid and carbohydrate [32] compositions, respectively. The complexes for ultracentrifugation were prepared by size-exclusion HPLC on Superose 6.

Crystals of the complex between NC10 scFv and tern N9 sialidase were grown as reported [14]. The crystals belong to the tetragonal space group $P42_12$ with unit-cell lengths $a = b = 14.10$ nm, $c = 21.79$ nm with two scFv/sialidase monomer complexes/asymmetric unit [14]. X-ray diffraction data were collected using the Weissenberg camera on beamline 6A2 (wavelength 0.100 nm) at the Photon Factory, Tsukuba, Japan [33]. The data were processed using the program WEIS [34] adapted to run on a Silicon Graphics Iris 4D/240VGX work station with display software IMAGE (Lawrence, M. C., unpublished data). Reflections with intensity, $I > 2\,\sigma(I)$ and a maximum resolution of 0.30 nm were scaled and merged in the PROTEIN system [35] with an overall $R_{merg}$ on intensities of 10.1% for the data set with 89% completeness to 0.30 nm.

The structure was solved using molecular-replacement techniques in the program XPLOR Version 3.0 [36] implemented on a Convex C3220 computer. The sialidase structure from the whale N9 sialidase/NC10 Fab structure [26] was used as the initial search model for molecular replacement. After the N9 sialidase monomer had been rotated and translated to its two correct positions in the asymmetric unit, the Fv fragments from the whale N9 sialidase/NC10 Fab complex were rotated and translated in the same manner, thus generating the complete asymmetric unit. The coordinates of tern N9 sialidase mutant Ser370→Leu [37] (Brookhaven Data Bank [38] entry 2NN9) were then substituted for the whale N9 sialidase coordinates. The appropriate mutations in each protein chain were made: in sialidase, Leu370→Ser; in $V_H$, Pro7→Ser, Leu109→Val, and Ser113→Gly; in $V_L$, Gln3→Glu, Met4→Leu, Arg108→Asp and Ala109→Tyr. The structure was then subjected to two rounds of crystallographic refinement in XPLOR interpolated with manual model building to correct the positioning of mutated side chains. The structure and electron-density maps were displayed with TURBO-FRODO [39] running on a Silicon Graphics Iris 4D/240VGX work station.

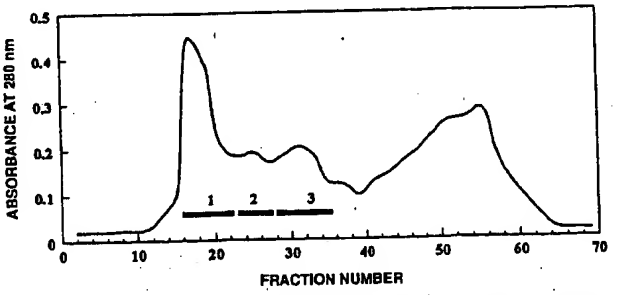



**Fig. 1. Sephadex G-100 gel filtration of the solubilized NC10 scFv (15-residue linker).** The column (60 cm×2.5 cm) was equilibrated with 0.025 M Tris/HCl, 0.1 M NaCl, pH 8.0, and run at a flow rate of 40 ml/h; 10-ml fractions were collected. Aliquots were taken across peaks 1–3 for SDS/PAGE analysis to locate the scFv using protein stain (see Fig. 2) and Western blot analysis. The three peaks were pooled as indicated by the bars.

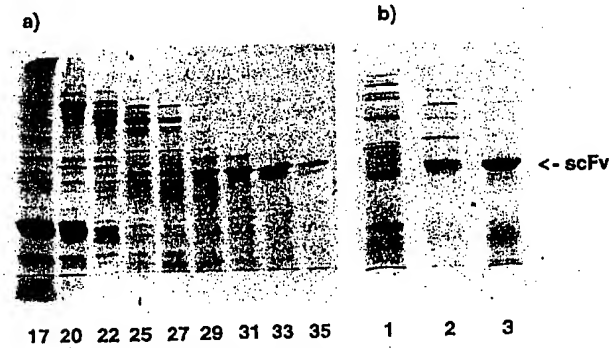



**Fig. 2. SDS/PAGE analysis of fractions from the Sephadex G-100 gel filtration.** (a) Fractions from peaks 1–3 in Fig. 1. Lanes from left to right are fractions 17–35 as indicated. (b) Pools 1–3. The SDS gels were stained with Coomassie brilliant blue G-250. Western blot analysis using the anti-FLAG M2 antibody confirmed that the major band at 27 kDa was scFv and that this was present mainly in peak 3 (data not shown).

## RESULTS

### scFv purification

Sephadex G-100 gel filtration of the soluble scFv fraction, obtained after extraction of the *E. coli* membrane fraction with 6 M guanidine hydrochloride and dialysis to remove the denaturant, resolved three protein peaks which were well separated from a large peak which contained pigmented, low-molecular-mass material (Fig. 1). SDS/PAGE analysis of fractions across peaks 1 to 3 showed that the scFv eluted mainly in peak 3 (Fig. 2a and b); Western blot analysis of the gel also detected scFv in peaks 1 and 2 but in lower amounts than in peak 3 (data not shown). Chromatography of pool 3 (Fig. 1) on a Mono-Q column (Fig. 3) yielded pure scFv as judged by SDS/PAGE and N-terminal analyses (data not shown). Gel filtration on a calibrated Superose 12 column showed that the purified scFv (Fig. 3) was a monomer eluting with an apparent molecular mass of approximately 30 kDa. SDS/PAGE yielded an apparent molecular mass of approximately 27 kDa. The purification data showed that on refolding from 6 M guanidine hydrochloride NC10 scFv formed predominantly monomer with only a small amount of dimer and higher-molecular-mass aggregates (Fig. 1). Gel filtration of pool 2 (Fig. 1) on Superose 12 confirmed the

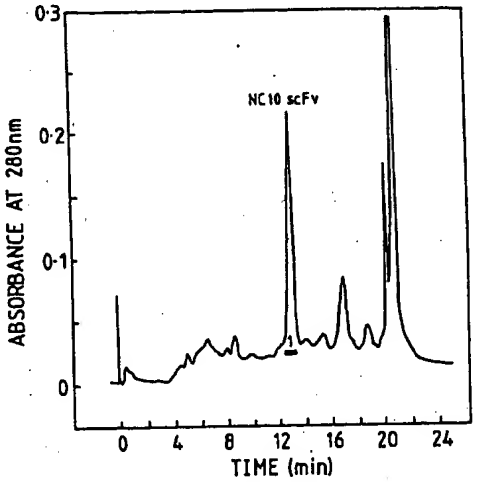



**Fig. 3. Chromatography of peak 3 (Fig. 1) on a Mono-Q HR5/5 column (Pharmacia-LKB).** The column was equilibrated with 25 mM Tris/HCl, pH 8.0 (buffer A) and a linear gradient to 30% 25 mM Tris/HCl, 1.0 M NaCl, pH 8.0 (buffer B) over 33 min was applied at a flow rate of 1.0 ml/min. The major peak (peak 1) containing scFv monomer was collected.

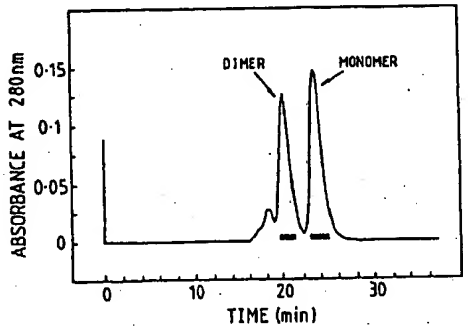



**Fig. 4. Size-exclusion HPLC on a calibrated Superdex 75 HR 10/30 column (Pharmacia-LKB) of purified NC10 scFv (peak 1, Fig. 3) after concentration and storage at −20°C for several weeks.** The column was equilibrated with 0.025 M Tris/HCl, 0.1 M NaCl, pH 8.0, and run at a flow rate of 0.5 ml/min. The monomer and dimer peaks were pooled as indicated for stability and affinity binding measurements.

presence of dimer in this pool. The dimer and higher-molecular-mass multimers (see also Figs 4 and 5) ran as monomers of approximately 27 kDa on non-reduced SDS gels which showed that the monomers are not linked by disulfide bonds.

### scFv monomer-dimer equilibrium

Gel filtration of purified scFv monomer (Fig. 3), which was concentrated to approximately 5 mg/ml for crystallization trials and stored frozen at −20°C for several weeks, revealed the presence of a mixture of monomer, dimer and higher-molecular-mass species (Fig. 4). The dimer eluted with an apparent molecular mass of approximately 54 kDa. The re-equilibration of these species was not rapid so they were readily separated by gel filtration of the concentrated scFv preparation. Rechromatography of the concentrated, purified scFv (Fig. 3), which contained the dimer and higher-

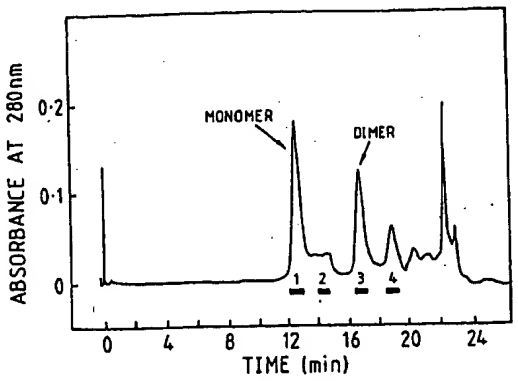



**Fig. 5. Rechromatography on Mono-Q of the purified scFv which contained dimer and higher-molecular-mass multimers.** Conditions were as described for Fig. 3.

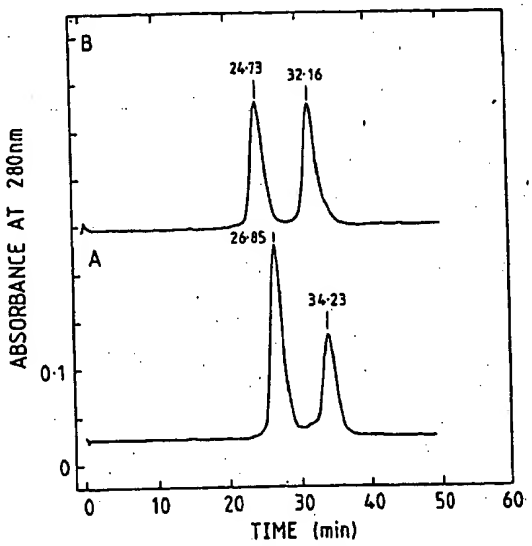



**Fig. 6. Size-exclusion HPLC on a calibrated Superose 6 HR10/30 column (Pharmacia-LKB) of the complex formed on the interaction of tern N9 sialidase with (a) scFv monomer and (b) scFv dimer.** In both cases the scFv was present in excess and the column was run under the same conditions as those in Fig. 4. The elution times are indicated for each peak.

molecular-mass species, on Mono-Q resolved at least five or six protein peaks (Fig. 5). Gel filtration showed that these major peaks corresponded to the monomer, dimer and higher-molecular-mass multimers.

The stability of monomer and dimer isolated by size-exclusion HPLC, and the conditions for the interconversion of monomer and dimer were studied. At a protein concentrations of 1.0 mg/ml in 0.025 M Tris/HCl, 0.1 M NaCl, pH 8.0, the monomer was stable for at least 17 days at 4°C and 20°C. At higher protein concentrations (≈ 5 mg/ml) the monomer aggregated to form dimer and higher-molecular-mass multimers, and repeated freezing and thawing increased the proportion of dimer and higher-molecular-mass multimers produced. The dimer at 1.0 mg/ml, however, dissociated to monomer; after 17 days at 20°C the dimer had completely dissociated to the monomer. At 4°C the dissociation of the dimer, at 1.0 mg/ml, was much slower with about 50% of the dimer remaining after 17 days. The dimer, at 1.0 mg/ml and at 20°C, did not dissociate in 1.0 M NaCl, 0.025 M Tris/HCl, pH 8.0, or conditions used to dissociate antibody-antigen binding (0.2 M glycine, pH 2.5); however, in the presence of 60% ethylene glycol, which suppresses hydrophobic interactions [40], the dimer and the higher-molecular-mass multimers dissociated to the monomer.

## Binding affinities of the monomer and dimer scFv fragments

Size-exclusion HPLC on a calibrated Superose 6 column showed that both the monomer and dimer of NC10 scFv interacted with tern N9 sialidase to form stable complexes (Fig. 6). The complex with the monomer eluted with an apparent molecular mass of approximately 300 kDa which is consistent with four monomers of scFv (each ≈ 27 kDa) binding/sialidase tetramer (190 kDa). The scFv dimer/sialidase complex, formed in the presence of excess scFv, eluted at an apparent molecular mass of approximately 520 kDa (Fig. 6); this is larger than the complex expected (407 kDa) if the dimer was only monovalent and smaller than the complex expected (596 kDa) if the dimer was bivalent and four dimers of scFv (each ≈ 54 kDa) cross-linked two sialidase tetramers. Since gel filtration behaviour depends on size and shape, sedimentation equilibrium experiments were run to define the molecular mass of the complexes. The molecular mass obtained for the scFv monomer/sialidase complex was 276 kDa in agreement with the gel-filtration result. The molecular mass of the scFv dimer/sialidase complex was 569 kDa. The size of this complex is consistent with four scFv dimers cross-linking two sialidase tetramers and shows that the dimer is bivalent. The linearity of the ln $c$ versus $r^2$ plots [41] of the sedimentation data showed that both complexes were homogeneous with respect to molecular mass (Fig. 7) and indicated that discrete, stoichiometric complexes were formed.

The binding affinities of both monomer and dimer scFv molecules for whale N9 sialidase were determined by sedimentation equilibrium [31]. The affinity constant ($K_a$) for the dimer ($2.6\times10^7\ M^{-1}$) was approximately two-fold higher than that for the monomer ($1.4\times10^7\ M^{-1}$). Analysis of the ultracentrifuge data showed that approximately 19% of the monomer and 7% of the dimer did not bind to the sialidase. This suggests that the $^{125}$I-radiolabelling of the scFv may have resulted in some loss of binding activity since gel-filtration studies of complex formation in the presence of excess sialidase indicated that both the monomer and dimer had full binding activity. A plot of ln $c$ versus $r^2$ of the sedimentation data [41] for the dimer gave a molecular mass of 54 kDa indicating that dimer had not dissociated to monomer to a significant extent during the measurement of the affinity constant in the airfuge.

## scFv/sialidase complex crystal structure

The scFv/sialidase structure has a crystallographic $R$ factor of 0.203 for data with $I > 2\ \sigma(I)$ in the resolution range 0.80–0.30 nm with root mean square (rms) $\Delta$ (bonds) of 0.0016 nm, rms $\Delta$ (angles) of 2.2°. Separate overall B-factors were allocated for each of the sialidase monomers and Fv fragments in the asymmetric unit; these refined to be 0.063 nm² and 0.061 nm², respectively, for sialidase monomers in complex 1 and complex 2, while the B-factors for the Fv in complex 1 and complex 2 refined to values of 0.0150 nm² and 0.0153 nm², respectively.

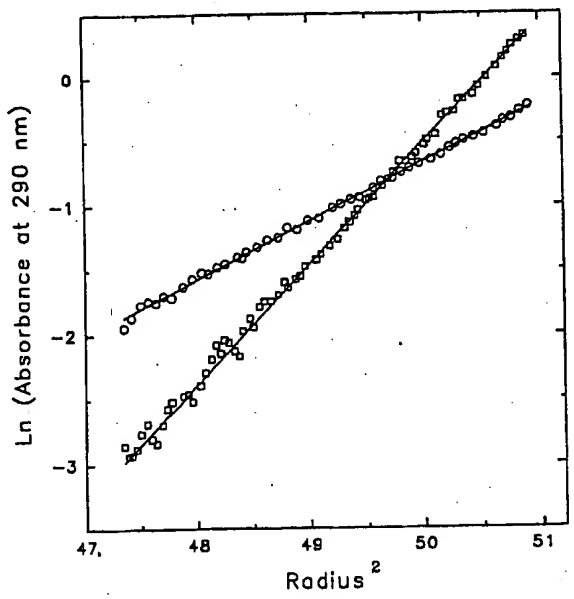



**Fig. 7. Sedimentation equilibrium data for complexes of tern sialidase and NC10 scFv monomer and dimer.** The complexes were isolated by size-exclusion chromatography (Fig. 6) in 0.05 M sodium phosphate, 0.15 M NaCl, pH 7.4. Experiments were conducted at 1960 *g* at 20°C for 24 h using a double sector centrepiece and 100 µl sample. The absorbance at 290 nm was determined as a function of radius in cm. (○), Data for the sialidase/scFv monomer; (□) data for the sialidase/scFv dimer.

The crystal structure showed two crystallographically distinct tetrameric scFv/sialidase complexes located on the crystallographic fourfold axis, *c*. The complexes are related by a translation of half a unit cell along *c* and a rotation of about 3° around the *c* axis. The two complexes assume the same basic conformation; they have a rms deviation of 0.040 nm over 1857 matched main-chain atoms. When only the main-chain atoms of the Fv fragments of the complexes are compared, the rms deviation is 0.047 nm, which indicates that there is a somewhat larger difference between the two scFv molecules in the asymmetric unit than between the two sialidase monomers. Crystal packing in the *ab* plane is facilitated by interactions between $V_H$ domains around the crystallographic diads. There are two distinct, but similar, interfaces involving symmetry pairs of each of the two scFv molecules in the asymmetric unit. One of these pairs is illustrated in Fig. 8.

The Fab/sialidase complex and the two independent scFv/sialidase complexes in the asymmetric unit were aligned by matching the 1800 identical main-chain atoms, which yielded a rms deviation of 0.067 nm for complex 1 and 0.069 nm for complex 2. Similar values of 0.067 (0.068) nm and 0.069 (0.070) nm, respectively, were achieved when the main-chain (C$\alpha$) coordinates of only the antibody variable domains of the complexes were compared.

## DISCUSSION

The NC10 scFv, in solution, upon refolding from 6 M guanidine hydrochloride, was predominantly a monomer of approximately 27 kDa and had full binding activity. The scFv monomer, at a protein concentration of 1 mg/ml or less, was stable at 4°C and 20°C and this molecule reacted with sialidase to yield a complex of 276 kDa which is consistent with four scFv monomers binding one sialidase tetramer.

The scFv monomer aggregated to form a dimer and higher-molecular-mass multimers at higher protein concentrations (5 mg/ml) and this process was enhanced on freezing and thawing. However, the isolated scFv dimer was unstable, dissociating to the monomer at both 4°C and 20°C with a half-life of approximately 8 days at 20°C. The higher-molecular-mass multimers and the dimer dissociated in 60% ethylene glycol suggesting that hydrophobic interactions [40] between the $V_H$ and $V_L$ domains of native monomeric scFv may be involved in the aggregation. The scFv dimer was active, and bound to sialidase to yield a complex of 569 kDa. The molecular mass of the dimer scFv/sialidase complex showed that scFv dimer is bivalent and is consistent with four dimers cross-linking two sialidase tetramers to form the complex. Higher-molecular-mass complexes in which only one or two scFv dimers cross-link two or more sialidase tetramers were not observed with the dimer. Furthermore the sedimentation data indicated that the isolated complexes were homogeneous with respect to molecular mass and supports the concept that discrete complexes are formed in which four scFv monomers bind to a sialidase tetramer or four scFv dimers bind to cross-link two sialidase tetramers. The anomalous behaviour of the complex formed by four scFv dimers cross-linking two sialidase molecules on size-exclusion HPLC suggests that this complex may have a rather compact structure.

The binding affinities of the scFv monomer ($1.4\times10^7$ $M^{-1}$) and dimer ($2.6\times10^7$ $M^{-1}$) were comparable to that found for the interaction with NC10 Fab and whale sialidase ($2.0\times10^7$ $M^{-1}$) [31]. The slightly greater affinity of the scFv dimer for whale N9 sialidase may simply reflect increased avidity compared to the monomeric scFv. It should be noted that the binding affinity measurements were performed at low concentrations ($\approx 10^{-9}$ M scFv) while the concentrations used to form complexes in the gel-filtration experiments were significantly higher. It is thus possible that at low concentrations of scFv dimer, the cross-linked complex was not formed and the measured binding constant reflects the binding of the dimeric scFv to a single site on sialidase.

Although a number of scFv fragments for different target antigens have now been synthesized in bacteria and characterized there are only two reports of scFv fragments forming dimers and higher-molecular-mass multimers [20, 21]. Griffiths et al. [20] purified by affinity chromatography several soluble scFv fragments against a series of self-antigens selected from a phage antibody library and found that after concentration by ultrafiltration their preparations contained monomers, dimers and higher-molecular-mass multimers. The stability of the isolated dimers or conditions for their formation were not reported. Recently, Holliger et al. [21] reported the construction of a series of bivalent and bispecific scFv fragments. They found that scFv constructs (with 15-residue polypeptide linkers) of the anti-phenyloxazolone antibody NQ11.7.22 produced only monomers while the scFv of anti-lysozyme antibody D1.3 (with the same linker) formed a mixture of monomer and dimer. These results and those presented in this study indicate that different scFv fragments with 15-residue linkers show differing propensities to form dimers. The scFv of NQ11.7.22 may form dimers under

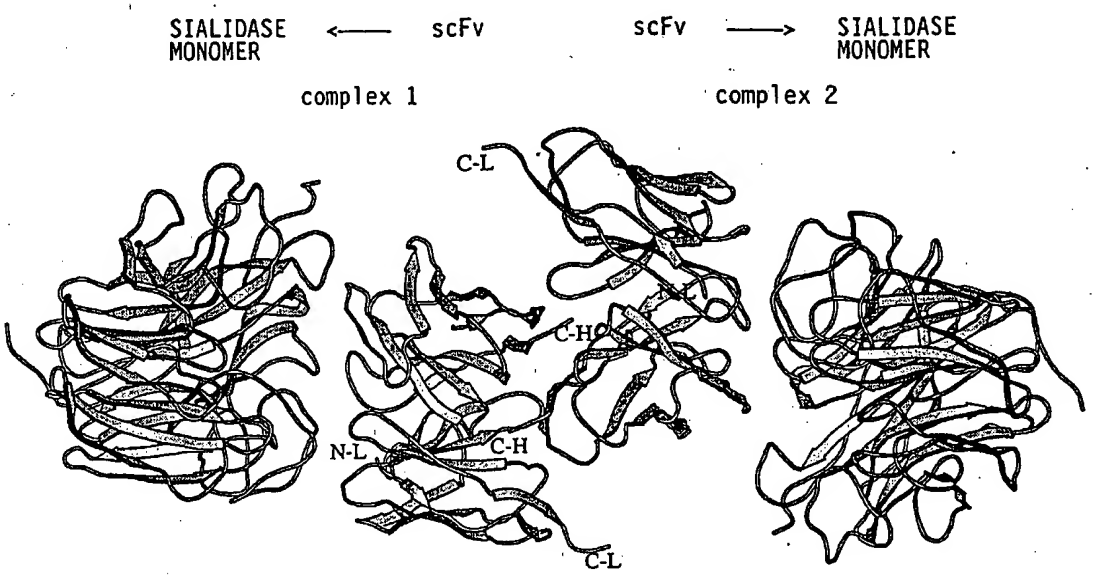



**Fig. 8. Structure of the NC10 scFv-tern N9 sialidase complex.** The diagram shows the close interaction between two complexs which are related by a crystallographic twofold rotation axis. The two scFv molecules are shown back to back and the binding interaction between the scFv and two different sialidase subunits is indicated by the arrows. The distance from the C-terminus of $V_H$ to the N-terminus of $V_L$ in complex 1 is 3.5 nm, while the distance from the C-terminus of the same $V_H$ to the N-terminus of the symmetry-related $V_L$ is only 1.8 nm. This figure was produced using MOLSCRIPT [43].

conditions similar to those described for the NC10 scFv whereas the D1.3 scFv [21] and other scFv [20] show a greater propensity to form dimers at low protein concentrations.

The solution of the crystal structure (Fig. 8) showed that both the mode of attachment and the binding interaction between the scFv and the sialidases are very similar to that in the parent Fab/sialidase complex. This is consistent with binding studies which showed comparable affinity for the scFv and Fab [28, 31]. When the structures of antibody D1.3 Fv and Fab in complex with lysozyme were compared [42] there was an rms difference between the positions of C$\alpha$ atoms in the antibody variable domains of 0.039 nm, compared with 0.069 nm in this study. It is likely that Fv and scFv will generally assume the same conformation as their parent Fab upon binding their target antigen.

At present there is no discernible electron density for the peptide linker which joins the $V_H$ and $V_L$ domains, which suggests that the linker is disordered in the crystal. This result was expected, given that the linker is mostly composed of glycine residues and that the predicted path of the linker leads around the exterior of the Fv and is solvent exposed. The distance from the C-terminus of the $V_H$ in complex 1 to the N-terminus of the $V_L$ to which it is paired to bind sialidase, is approximately 3.5 nm, compared to a distance of only 1.8 nm from the C-terminus of the same $V_H$ to the N-terminus of a $V_L$ which is related by a crystallographic twofold rotation (Fig. 8). A linker of 15 residues is long enough to span either distance comfortably; the possibility therefore arises that the scFv molecules have crystallized as 'crossed' dimers.

The solution of the scFv/sialidase complex crystal structure has demonstrated that scFv molecules can dock onto their target antigen in the same conformation as the parent Fab. One orientation by which dimeric scFv molecules can cross-link two sialidase subunits is illustrated in Fig. 8. The structure of scFv/sialidase complex suggests that it is possible for two target antigens to be cross-linked by interacting scFv molecules, oriented back-to-back across a twofold axis. Such an arrangement is plausible, rather than demonstrated, due to the absence of electron density for the linker peptide. Solution studies have demonstrated that scFv monomers can interact to form dimers which can cross-link two antigen molecules. To form the dimers in solution, the orientation of scFv dimers is likely to be different to that in the crystal dimer of Fig. 8.

The question of how the scFv (15-residue linker) dimers are formed remains to be resolved. Three possibilities are apparent: dimers (and higher-molecular-mass multimers) may form, by scFv fragments interlocking via the flexible linkers joining the $V_H$ and $V_L$ domains, or by the $V_H$ domain of one scFv molecule associating with the $V_L$ of another molecule to form a 'crossed' dimer (this mechanism appears to operate in the case when shorter linkers are used [21]), or by protein-protein interaction (e.g. via hydrophobic interactions) between the folded domains. The result with ethylene glycol dissociation of the dimer and the observed interaction between the $V_H$ domains in the crystal structure support the second and third possibilities for the dimer formation of scFv with longer linkers. Experiments are in progress to establish the bivalent character of these NC10 scFv dimers and to express scFv with truncated linkers to produce cross-linked dimers and bispecific molecules.

The authors thank Dr J. McKimm-Breschkin, Ms D. Marshall and Mr R. E. Guthrie for providing the samples of sialidase used in this study and Mr J. E. Burns for technical assistance in the protein purification and characterization. RLM was supported by a CSIRO Postgraduate Award and RGW was supported in part by grant AI 08831 from the National Institutes of Allergy and Infectious Diseases.

## REFERENCES


1. Winter, G. & Milstein, C. (1991) *Nature 349*, 293–299.
2. Orlandi, R., Gussow, D. H., Jones, P. T. & Winter, G. (1989) *Proc. Natl Acad. Sci. USA 86*, 3833–3837.
3. Neuberger, M. S., Williams, G. T. & Fox, R. O. (1984) *Nature 312*, 604–608.
4. Riechmann, L., Foote, J. & Winter, G. (1988) *J. Mol. Biol. 203*, 825–828.
5. Skerra, A. & Pluckthun, A. (1988) *Science 240*, 1038–1041.
6. Better, M., Chang, C. P., Robinson, R. R. & Horowitz, A. H. (1988) *Science 240*, 1041–1043.
7. Sawyer, J. R. & Blattner, F. R. (1991) *Protein Eng. 4*, 947–953.
8. Glockshuber, P., Malia, M., Pfitzinger, I. & Pluckthun, A. (1990) *Biochemistry 29*, 1362–1367.
9. Huston, J. S., Levison, D., Mudgett-Hunter, M., Tai, M., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R. & Opperman, H. (1988) *Proc. Natl Acad. Sci. USA 85*, 5879–5883.
10. Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, T., Pope, S. H., Riordan, G. S. & Whitlow, M. (1988) *Science 242*, 423–426.
11. Milencic, D. E., Yokota, T., Filupa, D. R., Finkelman, M. A. T., Dodd, S. W., Wood, J. F., Whitlow, M., Snoy, P. & Schlom, J. (1991) *Cancer Res. 51*, 6363–6371.
12. Denzin, L. K. & Voss, E. W. (1992) *J. Biol. Chem. 267*, 8925–8931.
13. Skerra, A., Pfitzinger, I. & Pluckthun, A. (1991) *Bio-technology (NY) 9*, 273–278.
14. Malby, R. L., Caldwell, J. B., Gruen, L. C., Harley, V. R., Ivancic, N., Kortt, A. A., Lilley, G. G., Power, B. E., Webster, R. G., Colman, P. M. & Hudson, P. J. (1993) *Proteins Struct., Funct. Genet. 16*, 57–63.
15. Whitlow, M. & Filpula, D. (1991) *Methods: Companion Methods Enzymol. 2*, 97–105.
16. Huston, J. S., Mudgett-Hunter, M., Tai, M-S., McCartney, J., Warren, F., Haber, E. & Opperman, H. (1991) *Methods Enzymol. 203*, 46–88.
17. Chaudhary, V. K., Queen, C., Junghans, R. P., Waldmann, T. A., Fitzgerald, D. J. & Pastan, I. (1989) *Nature 399*, 394–397.
18. Pack, P. & Pluckthun, A. (1992) *Biochemistry 31*, 1579–1584.
19. Hopp, T. P., Prickett, K. S., Price, V. L., Libby, R. T., March, C. J., Cerreti, D. P., Urdal, D. L. & Conlon, P. J. (1988) *Biotechnology (NY) 6*, 1204–1210.
20. Griffiths, A. D., Malmqvist, M., Marks, J. D., Bye, J. M., Embleton, M. J., McCafferty, J., Baier, M., Holliger, K. P., Gorich, B. P., Hughes-Jones, N. C., Hoogenboom, H. & Winter, G. (1993) *EMBO J. 12*, 725–734.
21. Holliger, P., Prospero, T. & Winter, G. (1993) *Proc. Natl Acad. Sci. USA 90*, 6444–6448.
22. Downie, J. C., Hinshaw, V. S. & Laver, W. G. (1977) *Aust. J. Exp. Biol. Med. Sci. 55*, 635–643.
23. Hinshaw, V. S., Bean, W. J., Geraci, J., Fiorelli, P., Early, G. & Webster, R. G. (1986) *J. Virol. 58*, 655–656.
24. Colman P. M., Laver, W. G., Varghese, J. N., Baker, A. T., Tulloch, P. A., Air, G. M. & Webster, R. G. (1987) *Nature 326*, 358–363
25. Tulip, W. R., Varghese, J. N., Laver, W. G., Webster, R. G. & Colman, P. M. (1992) *J. Mol. Biol. 227*, 122–148.
26. Colman, P. M., Tulip, W. R., Varghese, J. N., Tulloch, P. A., Baker, A. T., Laver, W. G., Air, G. M. & Webster, R. G. (1989) *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 323*, 511–518.
27. Air, G. M., Webster, R. G., Colman, P. M. & Laver, W. G. (1987) *Virology 160*, 346–354.
28. Gruen, L. C., Kortt, A. A. & Nice, E. (1993) *Eur. J. Biochem. 217*, 319–325.
29. Power, B. E., Ivancic, N., Harley, V. R., Webster, R. G., Kortt, A. A., Irving, R. A. & Hudson. P. J. (1992) *Gene (Amst.) 113*, 95–99.
30. McKimm-Breschkin, J. L., Caldwell, J. B., Guthrie, R. E. & Kortt, A. A. (1991) *J. Virol. Methods 32*, 121–124.
31. Gruen. L. C., McInerney, T. L., Webster, R. G. & Jackson, D. C. (1993) *J. Protein Chem. 12*, 255–259.
32. Ward, C. W., Murray, J. M., Roxburgh, C. M. & Jackson, D. C. (1983) *Virology 126*, 370–375.
33. Sakabe, N. (1991) *Nucl. Instrum. Methods A 303*, 448–463.
34. Higashi, T. (1989) *J. Appl. Crystallogr. 22*, 9–18.
35. Steigemann, W. (1974) PhD thesis. Technical University, Munich, Germany.
36. Brunger, A. T. (1992) *XPLOR reference manual version 3.0*, Molecular Simulations Inc., Waltham MA.
37. Tulip, W. R., Varghese, J. N., Baker, A. T., van Donkelaar, A., Laver, W. G., Webster, R. G. & Colman, P. M. (1991) *J. Mol. Biol. 221*, 487–497.
38. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.
39. Roussel, A. & Cambillau, C. (1989) in *Silicon graphics geometry partners directory, fall, Silicon Graphics eds*, pp. 77–78, Silicon Graphics, Mountain View CA.
40. Li, J. P. & Hjerten, S. (1991) *J. Biochem. Biophys. Methods 22*, 311–320.
41. Van Holde, K. E. (1975) in *The proteins*, (Neurath, H. & Hill, R. L. eds) vol 1, pp. 225–291, Academic Press, New York.
42. Bhat, T. N., Bentley, G. A., Fischmann, T. O., Boulot, G. & Poljak, R. J. (1990) *Nature 347*, 483–485.
43. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946–950.

The EMBO Journal vol.12 no.2 pp.725–734, 1993

# Human anti-self antibodies with high specificity from phage display libraries



Andrew D.Griffiths[1], Magnus Malmqvist[1], James D.Marks[1], Jacqueline M.Bye[1,2], M.J.Embleton[1], John McCafferty[3], Michael Baier[1], K.Philipp Holliger[1], Barbara D.Gorick[2], Nevin C.Hughes-Jones[2], Hennie R.Hoogenboom[1] and Greg Winter[1,4,5]

[1]MRC Centre for Protein Engineering, [2]MRC Molecular Immunology Unit, MRC Centre, Hills Road, Cambridge CB2 2QH, UK and Division of Transfusion Medicine, University of Cambridge, Cambridge CB2 2PT, [3]Cambridge Antibody Technology Ltd, The Science Park, Melbourn, Cambridgeshire SG8 6EJ and [4]MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK
[5]Corresponding author

Communicated by G.Winter

**Recently we demonstrated that human antibody fragments with binding activities against foreign antigens can be isolated from repertoires of rearranged V-genes derived from the mRNA of peripheral blood lymphocytes (PBLs) from unimmunized humans. The heavy and light chain V-genes were shuffled at random and cloned for display as single-chain Fv (scFv) fragments on the surface of filamentous phage, and the fragments selected by binding of the phage to antigen. Here we show that from the same phage library we can make scFv fragments encoded by both unmutated and mutated V-genes, with high specificities of binding to human self-antigens. Several of the affinity purified scFv fragments were shown to be a mixture of monomers and dimers in solution by FPLC gel filtration and the binding kinetics of the dimers were determined using surface plasmon resonance ($k_{on} = 10^5-10^6\ M^{-1}\ s^{-1}$, $k_{off} = 10^{-2}\ s^{-1}$ and $K_a = 10^7\ M^{-1}$). The kinetics of association are typical of known Ab–protein interactions, but the kinetics of dissociation are relatively fast. For therapeutic application, the binding affinities of such antibodies could be improved *in vitro* by mutation and selection for slower dissociation kinetics.**

*Key words*: human antibodies/phage display/self

## Introduction

Human monoclonal antibodies (mAbs) have huge potential for therapy, but are difficult to make by immortalizing B-lymphocytes (for reviews see James and Bell, 1987; Winter and Milstein, 1991). Furthermore, it is especially difficult to generate human mAbs directed against human antigens (anti-self antibodies), for example antibodies against soluble TNF to block septic shock (Spooner *et al.*, 1992), against membrane bound carcinoembryonic antigen to image colorectal carcinoma (Mach *et al.*, 1980) or against lymphocyte antigens to destroy tumour in lymphoma (Hale *et al.*, 1988). This difficulty results from immunological tolerance mechanisms that prevent the antigen-driven expansion of B-cell clones with self specificities (Burnet, 1959; Nossal, 1989). After antibody gene rearrangement, virgin B-cells may display antibodies with self-reactivity, but tolerance mechanisms can lead to their deletion (Nossal, 1989; Nemazee *et al.*, 1991; Russell *et al.*, 1991) or to their anergy (Nossal, 1989; Basten *et al.*, 1991; Erikson *et al.*, 1991). It has been suggested that cells may be anergized if the antigen is soluble, but deleted if the antigen is membrane bound (Hartley *et al.*, 1991). B-cell tolerance does not seem to occur when concentrations of soluble antigen are low (in contrast to T-cell tolerance) and B-cells with poor affinities for antigen are not tolerized, even at higher antigen concentrations (Adelstein *et al.*, 1991). Such non-tolerized B-cells are not usually expanded because they lack T-cell help (Bretscher and Cohn, 1970; Adelstein *et al.*, 1991), although proliferation can be induced artificially by using polyclonal B-cell activators (reviewed in Nossal, 1987).

It is estimated that 10–30% of B-lymphocytes in normal, healthy individuals are engaged in making autoantibodies (Cohen and Cooke, 1986). However, the 'natural autoantibodies' produced do not lend themselves to therapeutic use as they are often IgM, low affinity and polyreactive (see Nakamura *et al.*, 1988; Tomer and Schoenfeld, 1988; Casali and Notkins, 1989; Rossi *et al.*, 1990; Avrameas, 1991). An immune response against self can arise in autoimmune disease (see Smith and Steinberg, 1983) or after infections (see Bona, 1988) and a few human mAbs directed against self-antigens have been isolated from patients with active autoimmune disease (see James and Bell, 1987). These autoantibodies are frequently specific, but may bind to only a restricted range of epitopes on the antigen (see Bouanani *et al.*, 1991).

Recently monoclonal antibody fragments have been generated and expressed in bacteria using phage antibody technology (McCafferty *et al.*, 1990) by cloning repertoires of V-genes into filamentous bacteriophage and selecting the recombinant phage with antigen (for review, see Hoogenboom *et al.*, 1992). The repertoires comprised random combinatorial libraries (Huse *et al.*, 1989) of the rearranged heavy and light chain V-genes of immunized animals or human donors. Immunization leads to clonal expansion and production of mRNA by plasma cells: as a result, derived V-gene repertoires are enriched for sequences of heavy and light chains encoding part of an antigen binding site (Hawkins and Winter, 1992). The selected antibody fragments can have good affinities for antigen, for example at least $10^8\ M^{-1}$ for the hapten phOx (Clackson *et al.*, 1991). However, because it is difficult to raise an immune response to self-antigens, we have sought to extend the technology to the generation of human antibodies without the use of immunization.

In principle, a range of binding specificities could be isolated from a single huge and diverse phage library by selection with either self or foreign antigens (for review, see

Table I. Frequency of binding clones isolated from the unimmunized scFv library after selection

| Antigen | Rounds of selection | | | | | Number of unique clones |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| Thyroglobulin (bovine) | – | – | 18/40 | – | – | 12 |
| Thyroglobulin (human): selected on bovine | – | – | 10/40 | – | – | 4 |
| Fog1 (human IgGlκ antibody) | – | – | – | 94/96 | – | 4 |
| TNFα (human) | – | 122/1920 | 83/192 | 92/96 | – | 7 |
| CEA (human) | | – | 0/96 | 1/96 | 2/96 | 1 |
| MUC1 (human): selected with peptide | – | – | – | 0/96 | 2/96 | 1 |
| rsCD4 (human) | – | – | – | – | 8/96 | 1 |

The ratios indicate the frequency of binding clones after each round of selection. Phagemids were rescued with M13ΔgIII helper phage, except for the CEA, MUC1 and rsCD4 selections, where VCS-M13 helper phage was used.

Marks *et al.*, 1992a). To this end, from a large phage antibody library, we first isolated antibody fragments with a high specificity of binding to foreign antigens (turkey lysozyme, bovine serum albumin and the hapten phenyloxazolone) (Marks *et al.*, 1991) and with affinities ($K_a$) in the range $10^6-10^7$ $M^{-1}$. Repertoires of rearranged heavy and light chain V-genes were provided by PCR amplification from the $\mu$, $\kappa$ and $\lambda$ mRNA of peripheral blood lymphocytes from unimmunized, healthy human donors. The V-genes were assembled (Clackson *et al.*, 1991) at random (Huse *et al.*, 1989) to encode repertoires of single-chain Fv (scFv) fragments (Bird *et al.*, 1988; Huston *et al.*, 1988). The fragments were displayed on the surface of the filamentous bacteriophage (McCafferty *et al.*, 1990) by fusion to the minor coat protein pIII (Smith, 1985), and phage encoding scFv fragments with binding activities were selected by binding of the phage to antigen. On infection of bacteria with the selected phage (Hoogenboom *et al.*, 1991), soluble scFv fragments produced from individual clones by secretion into the bacterial periplasm (Glockshuber *et al.*, 1990) were screened for binding activity. We now demonstrate the use of the same phage library to isolate antibody fragments with high specificity against self-antigens.

## Results

### *The selected human antibody fragments show high specificity against self-antigens*

The unimmunized library was subjected to affinity enrichment on a range of antigens (see Materials and methods and Table I). After 2–5 rounds of selection, *Escherichia coli* cells were infected with eluted phage and antibody fragments produced by individual clones were screened for binding by ELISA. Phage selected with the 20 amino acid MUC1 peptide (Price *et al.*, 1990), which corresponds to a repeated motif in human MUC1 mucin (tumour-associated polymorphic epithelial mucin or PEM) (Gendler *et al.*, 1988; Gum *et al.*, 1990), were screened for binding to human PEM and hence bind to both peptide and the protein. The V-genes of clones with binding activities were sequenced and between 1–12 different clones identified for each antigen (Table I). The appearance of only low numbers of clones binding to CEA, PEM and human recombinant soluble CD4 (rsCD4), even after several rounds of selection, may reflect the use of VCS-M13 as helper phage (instead of M13ΔgIII helper used for the other antigens). Populations of phage(mid) particles produced by rescue with M13ΔgIII (which cannot produce pIII) have higher average avidities than those produced by rescue with VCS-M13 (where the wild-type pIII encoded by the helper phage can compete with scFv–pIII fusions).

The scFv fragments were then screened for binding to a panel of other protein antigens and were found to be highly specific. This is illustrated in Figure 1 with the three clones with strongest ELISA signals for bovine thyroglobulin, human TNFα and the human mAb Fog-1, and in Figure 2 with the single clones with binding activity to human CEA, MUC1 and human rsCD4. However for a few clones with poor ELISA signals on the target antigen, we found signals with some of the other proteins of the panel (not shown).

### *The antibody fragments are derived from a range of unmutated and somatically mutated V-genes*

The sequences of several clones with self-specificity are given in Table II and contain both kappa and lambda light chains (Table III). Comparison with the sequences of the nearest germline V-gene segments indicates that several different families are used (VH1, 3, 4 and 5; Vκ1 and 4, Vλ1, 2 and 3). In a few cases the V-genes are completely germline, for example both the VH and Vλ genes of αThy-29. However, most of the V-genes have several differences from the nearest germline V-gene segments, both at the nucleotide and amino acid level (Table III), suggesting that they are derived from somatically mutated B-cells (Berek and Milstein, 1987). Some mutations may have arisen during the PCR amplification and assembly process, for example the VH-genes of αFOG1-G8 and αMUC1-1, and the Vκ-gene of αThy-33 probably arose from cross-overs between two V-genes during PCR amplification (Table III). Furthermore, large differences (for example the Vκ of αFOG1-H6, which differs by 36 nucleotides) may be due to the use of unknown V-gene segments. There is a striking homology in the CDR3 of the heavy chain between αTNF-A1 and αTNF-E1: the germline V-genes are different but the same JH segments are used and 11 out of 16 residues of CDR3 are identical. This suggests that both scFv fragments may bind to the same epitope of TNF.

### *The antibody fragments are directed to different epitopes on the same protein*

The scFv fragments directed against bovine thyroglobulin were screened for binding to human thyroglobulin, which differs by only six single amino acid residues in the protomer (Malthiéry and Lissitzky, 1987). Four of the twelve clones (including αThy-29) bound to human thyroglobulin, whereas

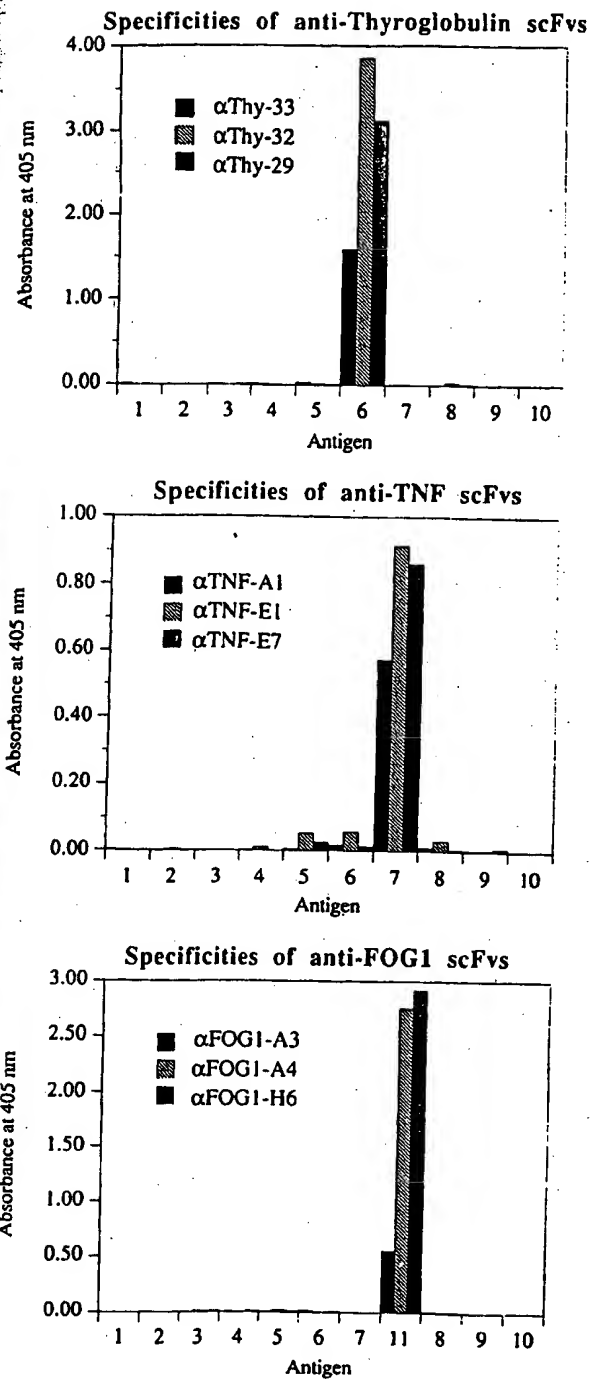



**Fig. 1.** Specificities of soluble single-chain Fvs (scFvs) isolated from the unimmunized library by selection on bovine thyroglobulin (upper panel), human TNFα (centre panel) or the human mAb Fog-1 ($\gamma 1,\kappa$) (Melamed *et al.*, 1987) (lower panel). Binding was determined by ELISA to a panel of proteins: 1, plastic; 2, hen egg trypsin inhibitor; 3, chymotrypsinogen A; 4, hen egg ovalbumin; 5, keyhole limpet haemocyanin; 6, bovine thyroglobulin; 7, human TNFα; 8, turkey egg-white lysozyme; 9, horse heart cytochrome c; 10, bovine serum albumin; 11, mAb Fog-1.

the rest (including αThy-32 and αThy-33) did not (data not shown). Likewise the fragments binding to the human antibody Fog-1 were screened for binding to a range of other antibodies differing in heavy and light chain isotype (Figure 3). The fragment αFOG1-A4 bound to all heavy

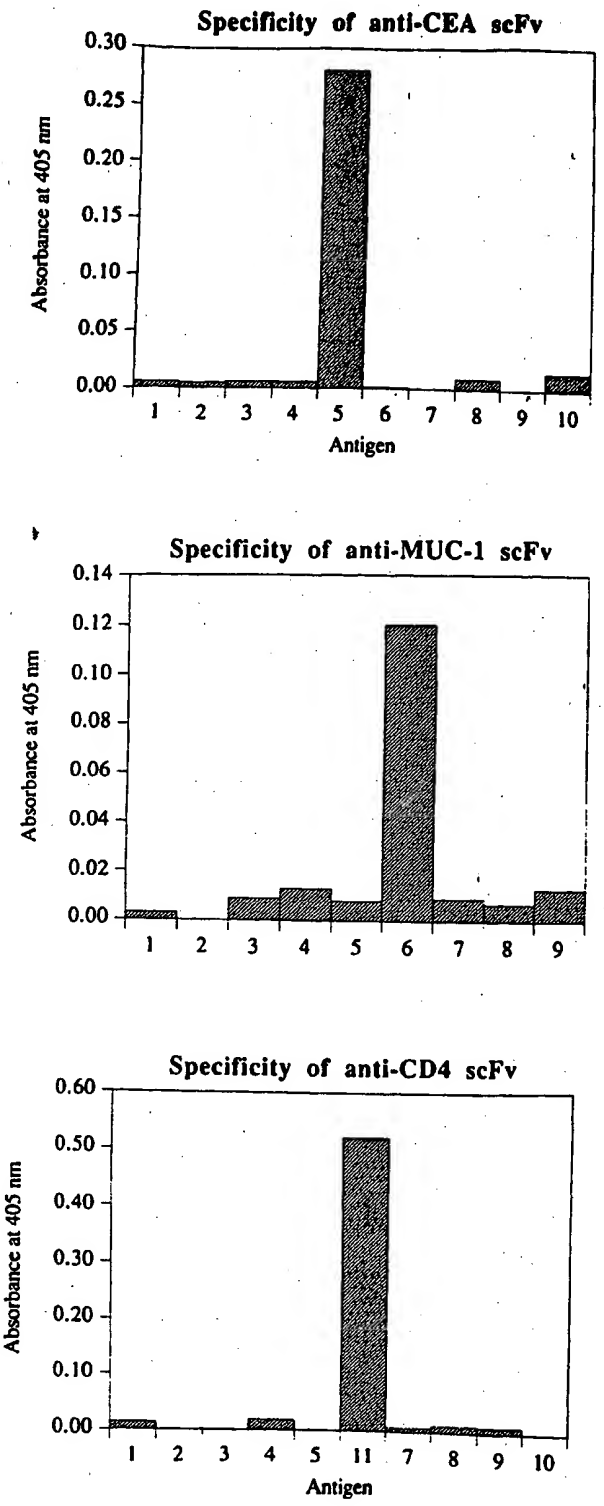



**Fig. 2.** Specificities of soluble single-chain Fvs (scFvs) isolated from the unimmunized library by selection on human CEA (upper panel), the MUC1 peptide (Price *et al.*, 1990) (central panel) or human CD4 (lower panel). Binding was determined by ELISA to a panel of proteins: 1, hen egg trypsin inhibitor; 2, chymotrypsinogen A; 3, hen egg ovalbumin; 4, keyhole limpet haemocyanin; 5, CEA; 6, urine extract containing human polymorphic epithelial mucin (PEM); 7, bovine thyroglobulin; 8, hen egg-white lysozyme; 9, bovine serum albumin; 10, chicken γ globulin coupled to 4-hydroxy-3-nitrophenyl acetic acid; 11, human recombinant soluble CD4.

**Table II.** Deduced protein sequences of several antigen-specific scFv fragments isolated from the unimmunized library

A. *Heavy chains*

| scFv | FR 1 | CDR 1 | FR 2 | CDR 2 | FR 3 | CDR 3 | FR 4 |
|---|---|---|---|---|---|---|---|
| αThy-23 | QVQLQQSGGGVVQPGRSMRLSCAASGFNFR | SYGMH | WVRQAPGKGLEWVS | GISGSGGSTYYADSVKG | RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAK | GSMIVVARYFDY | WGQGTLVTVSS |
| αThy-29 | QVQLVQSGAEVKKPGASVKVSCKASGYTFT | SYGIS | WVRQAPGQGLEWMG | WISAYNGNTNYAQKLQG | RVTMTTDTSTSTAYMELRSLRSDDTAVYYCAA | DTGRIDDFWSGYNFDY | WGQGTLVT |
| αThy-32 | QVQLVQSGGGLVQPGRSLRLSCAASGFTFD | DYAMH | WVRQAPGKGLEWVS | GISSNSGSIGYADSVKG | RFTISRDNAKNSLYLQMNSLRAEDTALYYCAR | GIAVAGAYYFDY | WGQGTLVTVSS |
| αThy-33 | QVQLVQSGGGVVQPGGSLRLSCAASGLSIR | TNGMH | WVRQAPGKGLEWVA | AISYDGRSVYYADSVQG | RVTISRDNSKNTVHLQITSLRSEDTAVYYCAK | DSSSSWFLDS | WGQGTLV |
| αThy-33 | QVQLVESGGGLVQPGGSLRLSCAASGFTFS | SYMMS | WVRQAPGKGLEWVA | NIKQDGSEKYYVDSVKG | RFTISRDNAKNTLYLQMNSLRDEDTAVYYCAK | NPRDSGSYYYFDY | WGQ |
| αFOG1-A3 | EVQLVESGGGLVQPGGSLRLSCAASGFTFS | NYAIH | WVRQAPGKGLEWVA | VISYDGGYEYYADSVKG | RSTISRDNSKNTLYLQMNSLRAEDTAVYYCAR | DASVHTAPYYYMDV | WGKG |
| αFOG1-A4 | QVQLQESGGSVVQPGRSLRLSCAASGFTFS | NSGMH | WVRQAPGKGLEWVS | YISSSSSTIYYADSVKG | RFTISRDNAKNSLYLQMNSLRDEDTAVYYCAR | EEGGLMDV | WGQ |
| αFOG1-H6 | QVQLQQSGSELVQPGGSLRLSCAASGFTFS | GYYMIG | WVRQMPGKGLEWMG | IINPGDSDTRYSPSFQG | QVTISVDKSVSTAYLQWSSLKPSDSAVYYCAR | HDVGYCSSPNCAKRPEYFQH | WGQ |
| αFOG1-G8 | QVQLQESGAGLLKPSETLSLTCAVYGGSFS | GYYWIG | WVLQAPGKGLEWVA | FIRYDGSNKYYADSVKG | RFTISRDNSKNTLYLQMNSLRAEDTAVYYCVR | EDHVITTGRYHYYMDV | WGK |
| αTNF-A1 | EVQLVESGGGLVQPGGSLRLSCAASGFTFS | SYGMH | WVRQAPGKGLEWVA | VISYDGSNKYYADSVKG | RFTISRDNSKNTLYLQMDNLRAEDTAVYYCAR | GPLRGYDYYYYYMDV | WGK |
| αTNF-E1 | QVQLQESGGGLVQPGGSLRLSCAASGLTFS | SYAMH | WVRQAPGKGLEWVA | VISYDGSNKYYADSVKG | RVTITADESTSTAYMELSSLRSEDTAVYYCAR | SGDLYSGYED | WGQ |
| αTNF-E7 | QVQLQESGAEVKKPGSSVKVSCKASGGTFS | SYAIS | WVRQAPGQGLEWMG | GIIPIFGTANYAQKFQG | RVTITRDRSMSTAYMELSSLRSEDTAMYYCAR | SGDLYSGYED | WGQ |
| αTNF-H9G1 | QVQLVQSGAEVKKTGSSVKVSCKTSGYTFT | YRYLH | WVRQAPGQALEWMG | WITPFNGNTNYAQKFQD | RVTITRDTSASTAYMELSSLRSEDTAVYYCAR | DFLSGYLDY | WGQGTLVTVSS |
| αCEA4-8A | QVQLQQSGAEVKKPGASVKVSCKASGYTFT | GYYMH | WVRQAPGQGLEWMG | WINPNSGGTNYAQKFQG | RVTITRDTSASTAYMELSSLRSDDTAVYYCAR | VSGYCSSTSSYDYYYYYMDV | WG |
| αMUC1-1 | QVQLVQSGAEVKKPGASVKVSCKASGYTFT | TYWIG | WVRQMPGKGLEYMG | LIYPGDSDTRYSPSFQG | QVTISADKSISTAYLQWSSLKASDTAMYYCAR | VSGYCSSTSSYDYYYYYMDV | WG |
| αCD4-74 | QVQLQQSGAEMKKPGESLKISCKGFGYDFS | TYWIG | WVRQMPGKGLEYMG | LIYPGDSDTKYSPSFQG | QVTISADKSISTAYLQWSSLKASDTAMYYCAR | VSGYCSSTSSYDYYYYYMDV | WGQGTLVTVSS |

B. *Light chains*

| scFv | FR 1 | CDR 1 | FR 2 | CDR 2 | FR 3 | CDR 3 | FR 4 |
|---|---|---|---|---|---|---|---|
| αThy-23 | DIQMTQSPSSLSASVGDSVTITC | QASQGIRNDLA | WYQQKPGKAPKLLIY | AASTLQS | GVPSRFSGSGSGTEFTLTINGLQPEDFATYYC | QQLGAYPLT | FGGGTKLEIKR |
| αThy-29 | SSELTQDPAVSVALGQTVRITC | QGDSLRSYYAS | WYQQKPGQAPVLVIY | GKNNRPS | GIPDRFSGSSSGNTASLTITGAQAEDEADYYC | NSRDSSGNLYVV | FGGGTKLTVLG |
| αThy-32 | QSVLTQPPSVSGAPGQRVTISC | TGSSNIGAGYDVH | WYQQLPGAAPKLLIY | GNSNRPS | GVPDRFSGSKSGTSASLAITGLQAEDEADYYC | QSYDSSLSGWV | FGGGTKVTVLG |
| αThy-33 | DVVMTQSPSTVSASVGDRVTLTC | RASQSISRWLA | WYQQKPGQPPKLLIY | WASTRKS | GVPDRFSGSGSGTDFTLTISSLQADDFATYYC | QHYDSFSPS | FGQGTKVEIK |
| αFOG1-A3 | QSALTQPASVSGSPGQSITISC | TGTSSDVGGYNYVS | WYQQYPGKAPKLLIY | EVSKRPS | GVSNRFSGSKSGNTASLTISGLQTEDEADYFC | SAYAPTGIMM | FGGGTKLTVLG |
| αFOG1-A4 | DIQMTQSPSSLSASVGDRVTITC | RASQGISNYLA | WYQQKPGKVPKLLIY | AASTLQS | GVPSRFSGSGSGTDFTLTISSLQPEDVAVYYC | QQYYSTPT | FGQGTKVEIKR |
| αFOG1-H6 | DIQMTQSPSTLSASIGDRVTITC | RASQNIDSWLA | WFQHKPGKAPKPLIY | GAFTLQN | GVPSRFSGSGSGTEFSLTISSLQPDDFATYFC | QQAHSFPPT | FGGGTRLEIKR |
| αFOG1-G8 | DIQMTQSPSTLSASIGDRVTITC | RASQSIGLWLA | WYQQKPGKAPKLLIY | AASTLQT | GVPSTFSGSGSGTEFTLTISSLRPEDFATYYC | QQLISYPLT | FGGGTKVEIKR |
| αTNF-A1 | DIQMTQSPSSLSASVGDRVTITC | RESQGIRNDLG | WYQQKPGKAPKLLIY | GTSSLQS | GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC | QQTTSFPLT | FGGGTKLEIK |
| αTNF-E1 | EIVLTQSPSSVSASVGDRVTITC | RASQGIRSWLA | WYQQKPGKAPNLLIY | AASTLET | GVPSRFSGSGSGTDFTLTISSLQPEDFATYFC | QQANSFPLT | FGGGTKLE |
| αTNF-E7 | DIVMTQSPSSLSASIGDRVTITC | RESQGIRNDLG | WYQQKPGKAPKLLIY | AASSLQS | GVPSRFSGSGSGTDFTLTISSLQPEDFATYFC | QQANSLPPT | FGGGTKVEIKR |
| αTNF-H9G1 | QSVLTQPPSVSGPPGQRVTISC | TGTSSNIGAGYDVY | WYQQLPDTSPRVLIY | RNSNRPS | GVPDRFSGSKSGTSASLAITGLQAEDEADYYC | QSYDSRLIRV | FGGGTKLTVLG |
| αCEA4-8A | EIVLTQSPSSLSASVGDRVTITC | RASQSISSYLN | WYQQKPGKAPKLLIY | AASSLQS | GVPSRFSGSGSGTDFTLTISSLQPEDFATYYC | QQYSNYPLT | FGGGTKVDIK |
| αMUC1-1 | QSVLTQPASVSGSPDQSITISC | TGTSSDVGGYKYVS | WYQQHPDKAPKVMIY | DVTNRPS | GGSNRFSGSKSGNTASLTISGLQAEDEADYYC | SSYAGAQSLV | FGGGTKLTVLG |
| αCD4-74 | HVILTQPPSASVAPGQRVTISC | SGSRSNIGSNPVS | WYRQFPGAAPKLLIY | IDDRRSS | GVPDRVSGSRSGTSASLAISGLQSEDEADYYC | VAWDDSLKGHV | FGEGTKLTVL |

FR, framework region; CDR, complementarity-determining region. Bovine thyroglobulin binders; αThy-23, αThy-29, αThy-32 and αThy-33. Human thyroglobulin binders; αThy-23 and αThy-29. Fog-1 (a human γ1/κ mAb) binders: αFOG1-A3, αFOG1-A4, αFOG1-H6 and αFOG1-G8. Human TNFα binders; αTNF-A1, αTNF-E1, αTNF-E7 and αTNF-H9G1. Human CEA binder; αCEA4-8A. Human MUC1-1 binder; αMUC1-1. Human rsCD4 binder; αCD4-74.

chain γ1, 2 and 3 isotypes, but not to γ4 or μ. By contrast, the fragments αFOG1-H6 and αFOG1-A3 did not bind to any of the other antibodies, including those of the same isotype as Fog-1, suggesting that they are directed to the variable domain of Fog-1.

### *Two of the antibody fragments are directed against idiotopes of human mAb Fog-1*

The binding of $^{125}$I-Fog-1 antibody to human red blood cells bearing the Rh D antigen could be inhibited by both αFOG1-H6 and αFOG1-A3 scFv fragments. Hence, both αFOG1-H6 and αFOG1-A3 are site-associated anti-idiotype antibodies, complexing with the antigen-binding site of Fog-1. The extent of inhibition of $^{125}$I-Fog-1 binding to the Rh D antigen (on human $R_1R_2$ red blood cells) was determined by titration with affinity purified αFOG1-H6 and αFOG1-A3 scFv fragments. [As control, no inhibition of $^{125}$I-Fog-1 binding was observed using a scFv fragment (αTEL9) (Marks *et al.*, 1991) directed against turkey egg-white lysozyme.] With the maximum of 16 μg scFv (1000-fold molar excess to $^{125}$I-Fog-1), the binding was inhibited by 14.2% (αFOG1-H6) and 20.9% (αFOG1-A3), suggesting that the affinities of these fragments for Fog-1 are much lower than the affinity of Fog-1 for the Rh D antigen ($K_a = 2.2\times10^9$ M$^{-1}$) which binds monovalently (Gorick *et al.*, 1988). If 100% of the fragments are active, the affinities of the two fragments for binding to Fog-1 could be estimated as $K_a = 3\times10^5$ M$^{-1}$ for αFOG1-H6 and $6\times10^5$ M$^{-1}$ for αFOG1-A3 and this is consistent with other kinetic measurements (see below and Table IV).

### *The scFv fragments can form both monomers and dimers in solution*

Soluble antibody fragments were purified from bacterial supernatants by affinity chromatography, by binding of the C-terminal peptide tag to the mAb 9E10 (Munro and Pelham, 1986; Clackson *et al.*, 1991; Marks *et al.*, 1991). After ultrafiltration, the fragments were further purified by FPLC gel filtration (Pharmacia) on Superdex 75 (Pharmacia), and detected on-line both by UV absorption (280 nm) and by binding to antigen immobilized on a sensor chip in BIAcore (Pharmacia Biosensor AB) (Jönsson *et al.*, 1991; Jönsson and Malmqvist, 1992). This showed that the scFv fragments emerged in two peaks, corresponding in size to monomers and dimers (Figure 4). The dimers bind more strongly to the immobilized antigen than monomers due to their greater avidity of binding. The scFv dimers run as monomers on non-reducing SDS gels (Laemmli, 1970) (not shown) and are therefore not linked by disulphide bonds. As two peaks are seen in gel filtration, it appears that in this case the monomers and dimers do not interconvert rapidly (for discussion and references for gel filtration equilibria see Jones *et al.*, 1985). Presumably the dimers are scFv fragments interlocked through the flexible linker joining the heavy and light chains, or with the heavy chain of one scFv molecule associated with the light chain of the other. We have preliminary evidence that antibody Fab fragments made in bacteria can also multimerize (unpublished data).

### *The scFv fragments have micromolar affinities*

The presence of both scFv monomers and dimers could lead to an overestimate of affinity of binding using solid phase

Table III. V-gene family, germline derivation and extent of somatic hypermutation of several antigen-specific scFv fragments isolated from the unimmunized library

| scFv | Family | Germline genes of closest nucleotide sequence | Differences from germline | |
|---|---|---|---|---|
| | | | Nucleotide | Aminoacid |
| *Heavy chains* | | | | |
| αThy-23 | VH3 | DP-47 | 13 | 8 |
| αThy-29 | VH1 | DP-14 | 0 | 0 |
| αThy-32 | VH3 | DP-31 | 5 | 2 |
| αThy-33 | VH3 | DP-49 | 32 | 19 |
| αFOG1-A3 | VH3 | DP-54 | 7 | 3 |
| αFOG1-A4 | VH3 | DP-46 | 7 | 7 |
| αFOG1-H6 | VH3 | DP-51 | 10 | 4 |
| αFOG1-G8[a] | VH4 | DP-63 (FR1) | 2 | 0 |
| | VH5 | DP-73 (CDR1 to FR3) | 15 | 7 |
| αTNF-A1 | VH3 | DP-50 | 9 | 6 |
| αTNF-E1 | VH3 | DP-46 | 14 | 6 |
| αTNF-E7 | VH1 | DP-10 | 0 | 0 |
| αTNF-H9G1 | VH1 | DP-4 | 1 | 1 |
| αCEA4-8A | VH1 | DP14 | 1 | 0 |
| αMUC1-1[a] | VH1 | VI-2 (FR1 to CDR2) | 2 | 0 |
| | VH1 | DP-25 (FR3) | 0 | 0 |
| αCD4-74 | VH5 | DP-73 | 13 | 8 |
| *Light chains* | | | | |
| αThy-23 | Vκ1 | L8 | 20 | 9 |
| αThy-29 | Vλ3 | IGLV3S1 | 0 | 0 |
| αThy-32 | Vλ1 | IGLV1S2 | 1 | 1 |
| αThy-33[a] | Vκ1 | L12 (FR1 and CDR1) | 6 | 3 |
| | Vκ4 | B3 (FR2 to FR3) | 5 | 5 |
| αFOG1-A3 | Vλ2 | VL2.1 | 16 | 9 |
| αFOG1-A4 | Vκ1 | O4 | 25 | 12 |
| αFOG1-H6 | Vκ1 | L5 | 36 | 17 |
| αFOG1-G8 | Vκ1 | L8 | 25 | 10 |
| αTNF-A1 | Vκ1 | L11 | 12 | 8 |
| αTNF-E1 | Vκ1 | L5 | 5 | 5 |
| αTNF-E7 | Vκ1 | L11 | 17 | 8 |
| αTNF-H9G1 | Vλ1 | IGLV1S2 | 18 | 9 |
| αCEA4-8A | Vκ1 | O2 | 4 | 0 |
| αMUC1-1 | Vλ2 | VL2.1 | 18 | 12 |
| αCD4-74 | Vλ1 | Humlv1L1 | 23 | 17 |

References for all the heavy chain germline genes can be found in Tomlinson *et al.* (1992). The references for the light chains are VL2.1 (Brockly *et al.*, 1989); IGLV1S2 (Bernard *et al.*, 1990); IGLV3S1 (Frippiat *et al.*, 1990); L8(Vd) and L5(Vb)(Pech *et al.*, 1984); L12(HK102) (Bentley and Rabbits, 1980); B3(VKIV) (Klobeck *et al.*, 1985); 02 and 04 (Pargent *et al.*, 1991); L11 (Scott *et al.*, 1991); Humlv1L1 (Daley *et al.*, 1992). Alternative names are given in parenthesis. a) These genes appear to have been created by cross-overs between two V-genes during PCR amplification and therefore matches have been determined using the two putative germline segments: FR, framework; CDR, complementarity-determining region.

methods. To determine the affinity and kinetics of binding of scFv fragments to the antigen-coated chip using surface plasmon resonance, we therefore purified the fragments by gel filtration (Table IV). For the dimers, the off-rate constants were determined as $\sim 10^{-2}$ s$^{-1}$ and the on-rate constants for the scFv dimers as $\sim 10^{5}-10^{6}$ M$^{-1}$ s$^{-1}$ (assuming the sample is completely active). In the case of αFOG1-H6, the antigen (the mAb Fog-1) was immobilized on the sensor chip in two ways, either directly (Figure 5) or via a rabbit anti-mouse IgG1 antibody. The results were almost identical by either method (see Table IV). However the active fraction of scFv fragments varies considerably and could lead to an underestimate of the on-rate (and affinity of binding); for example using fluorescence quench titration with several scFv fragments directed against phenyloxazolone we detected only 0.06−0.38 functional binding sites per scFv molecule (unpublished data). Indeed the on-rate constants calculated for the association of the αFOG1-H6 fragment and Fog-1 antibody depend on whether the antibody ($k_{on}$ $2.2\times10^{5}$ M$^{-1}$ s$^{-1}$) or scFv fragment ($k_{on}$ $1.0\times10^{6}$ M$^{-1}$ s$^{-1}$) is immobilized on the sensor chip (Table IV), indicating that the αFOG1-H6 fragment is less active than the Fog-1 antibody. For the scFv monomers, the binding signals were low and it was difficult to follow the kinetics of binding to the surface, except for the dissociation of the αThy-29 monomer ($k_{off}$ = $2\times10^{-2}$ s$^{-1}$). However, the 4-fold stabilization of the αThy-29 fragment dimer (see below), suggests that the off-rate constants of the other monomers are $>10^{-2}$ s$^{-1}$, perhaps $10^{-1}$ s$^{-1}$.

The greater stability of the scFv dimers on the sensor chip

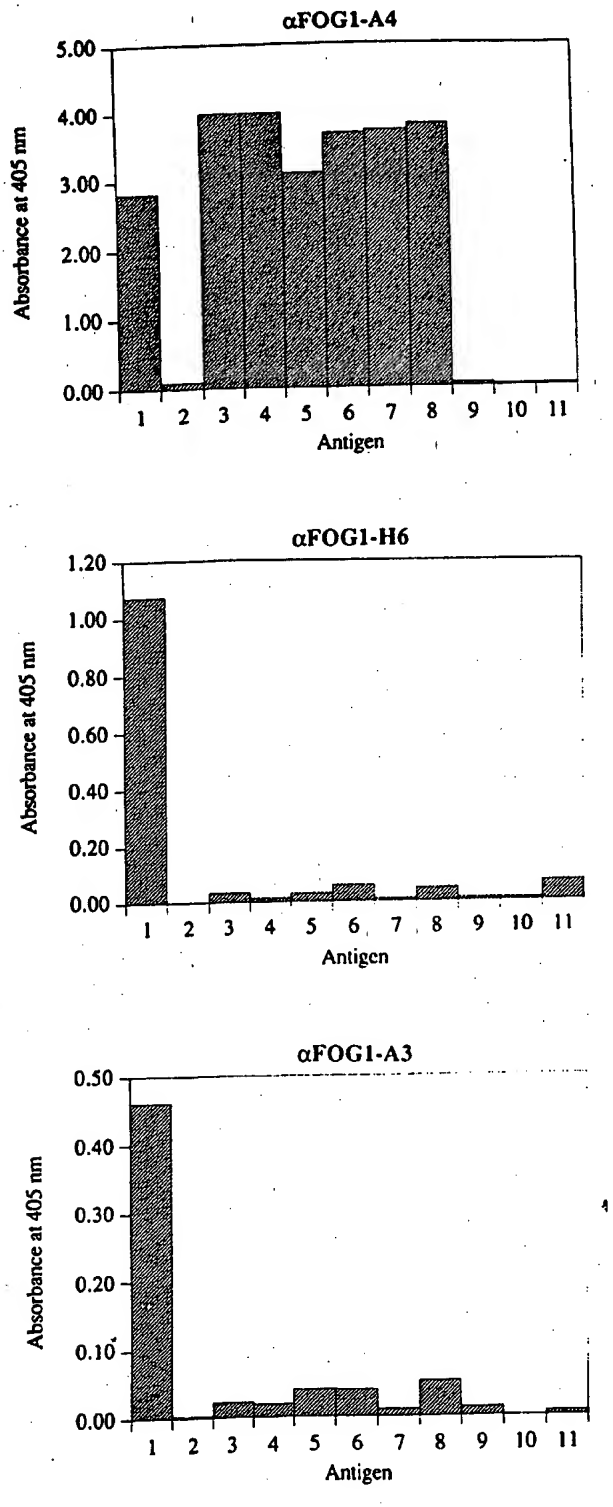



Fig. 3. ELISA to assay the binding of three scFvs, isolated by selection on a human monoclonal antibody Fog-1 (IgG1, κ) (Melamed *et al.*, 1987), to a panel of human antibodies of varying isotype: 1, Fog-1; 2, the Fv fragment of Hulys11 (Foote and Winter, 1992); 3, Hulys11 antibody (IgG1, κ); 4, RegA (IgG1, κ) (Melamed *et al.*, 1987); 5, FogC (IgG3, κ) (N.C.Hughes-Jones, unpublished); 6, Pag1 (IgG1, λ) (Thompson *et al.*, 1986); 7, IgG2, λ antibody purified from myeloma plasma (Sigma); 8, Oak3, (IgG3, λ) (Bye *et al.*, 1992); 9, IgG4, λ purified from myeloma plasma (Sigma); 10, Fom1 (IgM, λ) (Melamed *et al.*, 1987); 11, FomA (IgM, λ) (Melamed *et al.*, 1987).

compared with monomers indicates that the dimers are bivalent. The scFv dimers are therefore analogous to the two heads of the antibody IgG, but with different spacing between the heads, and their binding avidities were estimated as $\sim 10^7$ $M^{-1}$ from $k_{on}/k_{off}$ (Table IV). The affinities of the monomers must be lower by virtue of their faster dissociation from the surface. For the αThy-29 monomer, and assuming that the on-rate constant is the same as for the dimer (Mason and Williams, 1986), we can estimate an affinity of $\sim 3\times10^6$ $M^{-1}$. These affinities, calculated from the rate constants measured by surface plasmon resonance, appear to be similar to those measured in solution by fluorescence quench techniques. For example the affinity of binding of the monomer scFv fragment αTEL9 (Marks *et al.*, 1991), which binds to turkey lysozyme (and was derived from the same library), was estimated as $3.9\times10^7$ $M^{-1}$ using surface plasmon resonance (Table IV), and as $1.2\times10^7$ $M^{-1}$ by fluorescence quench (Marks *et al.*, 1991).

## Discussion

We had demonstrated previously that highly specific human antibody fragments (scFv), directed against 'foreign' antigens, both protein and hapten, could be isolated from a large phage display library composed of the rearranged V-genes of peripheral blood lymphocytes from unimmunized donors (Marks *et al.*, 1991). Here we have shown that antibody fragments directed against human antigens (self-antigens), including idiotopes of a human antibody, a cytokine (TNFα), two tumour markers (CEA and MUC1) and the T-lymphocyte marker CD4 can be derived from the same library. Immunological tolerance would make immunization of humans with these antigens difficult; TNFα is also extremely toxic and raising an immune response against CD4 would be an act of suicide by the immune system. The antibody fragments we have isolated show a high specificity of binding to antigen. This contrasts with the poor specificity of binding of fragments isolated from a phage display library in which mouse Fab fragments were fused to the major coat protein (pVIII) of filamentous phage (Gram *et al.*, 1992). The use of different V-genes and heavy and light chain combinations for each antigen suggested that each fragment was likely to bind to different epitopes and this was shown directly for the fragments against thyroglobulin and the human mAb Fog-1.

The affinity of antibodies isolated from a library is thought to be proportional to the library size (Perelson and Oster, 1979) and in this case the size of the phage library is comparable to the number of B-cells in a mouse, and the affinities of antibodies isolated are typical of antibodies from the mouse primary immune response (Foote and Milstein, 1991). The kinetics of association of the antibody fragments to the protein self-antigens ($10^5-10^6$ $M^{-1}$ $s^{-1}$) are also typical of previously characterized Ab−protein interactions. However the kinetics of dissociation ($10^{-2}$ $s^{-1}$) are relatively fast for Ab−protein interactions (but both rates are slow compared with many Ab−hapten interactions) (Smith and Skubitz, 1975; Pecht, 1982; Mason and Williams, 1986; Foote and Milstein, 1991; Foote and Winter, 1992). At first sight, it is surprising that we can isolate scFv fragments with such fast off-rates, as a 'monomeric' phage should not be retained on the solid support during washing. However, scFv fragments are

Table IV. Affinities and kinetics of antigen binding by monomeric and dimeric scFv fragments

| scFv | (M/D)[a] | Immobilized species | $k_{on}$[b] (BIAcore) $M^{-1}s^{-1}/10^4$ | $k_{off}$[b](BIAcore) $s^{-1}/10^{-2}$ | $K_a=k_{on}/k_{off}$ (BIAcore) $M^{-1}/10^6$ | $K_a$ by FQ[c] or inhibition[d] $M^{-1}/10^6$ |
|---|---|---|---|---|---|---|
| αTNF-E7 | D | HumanTNFα | 9.0 (±1.2) | 1.4 (±0.054) | 6.4 | ND |
| αFOG1-H6 | D | Fog-1 (direct) | 22.2 (±0.4) | 1.8 (±0.23) | 12.3 | ND |
| αFOG1-H6 | D | Fog-1 (via RAMIgG1) | 22.1 (±1.9) | 2.4 (±0.045) | 9.3 | ND |
| αFOG1-H6 | D | αFOG1-H6 scFv | 104 (±2.4) | ND[e] | ND | ND |
| αFOG1-H6 | M + D | (Measured by inhibition) | ND | ND | ND | 0.3[d] |
| αFOG1-A3 | M + D | (Measured by inhibition) | ND | ND | ND | 0.6[d] |
| αThy-29 | D | Human thyroglobulin | 6.6 (±1.2) | 0.46 (±0.063) | 14.3 | ND |
| αThy-29 | M | Human thyroglobulin | ND | 2.0 (±0.37) | ND | ND |
| αTEL9 | M | Turkey egg lysozyme | 39.2 (±2.6) | 1.0 (±0.97) | 39.2 | 11.6[c] |

[a] M, monomeric fraction; D, dimeric fraction.
[b] Numbers in brackets are standard deviations.
[c] FQ, fluorescence quench titration.
[d] Calculated from the extent of inhibition of $^{125}$I-Fog-1 binding to the Rh D antigen.
[e] Not determined because the dissociation curves were very badly bent.

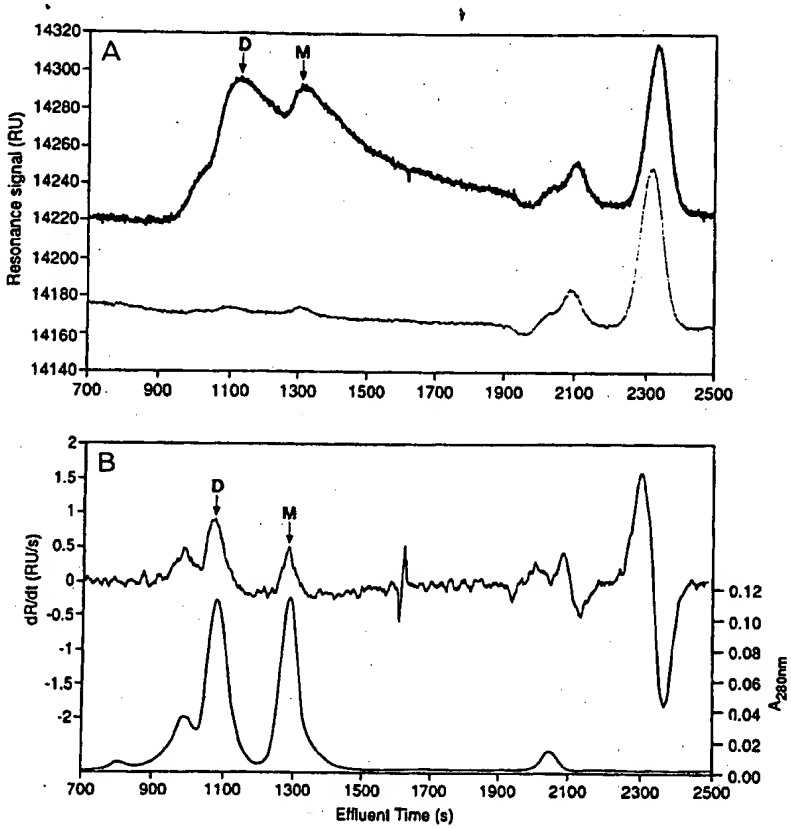



Fig. 4. Gel filtration of affinity purified soluble scFv αThy-29 on Superdex 75 analysed by UV absorption and on-line specific detection of the active component on BIAcore. A. BIAcore sensorgram [resonance signal (RU) as a function of time] showing adsorption of scFv in the column effluent passing over a sensor chip with immobilized human thyroglobulin (thick line) and the same sample run over a bare CM5 sensor chip surface without any antigen (thin line). B. UV profile of the gel filtration (lower line) and the derivatized sensorgram (upper line) which illustrates the rate of change in mass of protein bound to the sensor chip as a function of time. M, scFv monomer; D, scFv dimer.

displayed multivalently on the phage, especially using the M13ΔgIII helper phage, and some of the scFvs that tend to form dimers in solution may also form dimers on phage. The multivalent interactions with antigen help retain the phage, allowing the encoded scFv phage to be isolated.

Random combinatorial V-gene repertoires derived from the mRNA of immunized animals are enriched for heavy or light chain V-genes encoding part of an antigen binding site (Winter and Milstein, 1991; Hawkins and Winter, 1992), and this facilitates the isolation of antigen-binding fragments (Clackson *et al.*, 1991; Persson *et al.*, 1991) although the combinations of V-genes of each B-lymphocyte appear to be largely destroyed (Winter and Milstein, 1991; Gherardi and Milstein, 1992). Antigen binding sites can also be generated *de novo* by the random combination of chains, as illustrated by the isolation of scFv fragments against

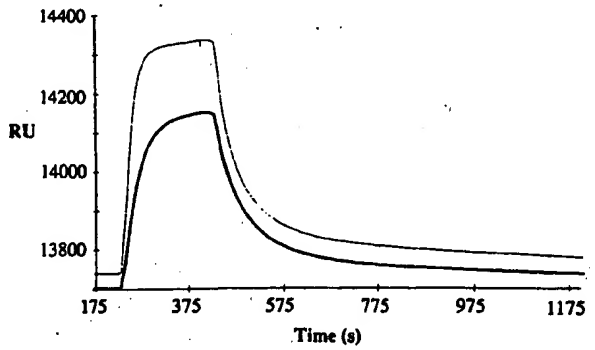



**Fig. 5.** BIAcore sensorgram [resonance signal (RU) as a function of time] of the interaction of soluble αFOG1-H6 scFv dimer with immobilized mAb Fog-1 (Melamed *et al.*, 1987). A 35 μl pulse of 200 nM (thin line) or 80 nM (thick line) scFv protein was passed, with a flow rate of 10 μl/min over a sensor chip to which mAb Fog-1 was coupled directly.

foreign antigens from unimmunized human donors (Marks *et al.*, 1991). However, the origins of the V-genes of scFv fragments directed against self-antigens are less clear. Self-reactive antibodies, including those with specificities against human thyroglobulin (Ruf *et al.*, 1985), human TNFα (Bendtzen *et al.*, 1990) and human IgG (Welch *et al.*, 1983), are common in healthy individuals and indeed 10–30% of B-lymphocytes appear to be engaged in making autoantibodies (Cohen and Cooke, 1986). Therefore the V-genes could be derived from B-cells that are autoreactive, or those that are not. Since somatic hypermutation of antibody genes is triggered only after antigen-induced B-cell proliferation (Griffiths *et al.*, 1984), the isolation of scFv fragments encoded by somatically mutated V-genes (Table III) indicates that the V-genes have been derived from lymphocytes that have been stimulated by antigen; for example from B-cells with self-specificities that have been stimulated with cross-reactive foreign antigen, or from B-cells encoding antibodies of other (foreign) specificities. Conversely those scFv fragments encoded by V-genes with little or no somatic mutation (see Table III) may well have been derived from virgin B-cells or those involved in early immune responses.

'Natural autoantibodies' (self-reactive antibodies isolated from healthy donors) tend to be of low affinity and polyspecific and may well be produced by a discrete subset of B-cells, the internal activity set (Holmberg and Coutinho, 1985), contributed in part by $CD5^+$ B-cells (Casali and Notkins, 1989). In contrast, these anti-self scFv fragments are highly specific in binding to antigen despite only having micromolar affinities. However, their affinities could presumably be improved *in vitro*, for example, the affinity of an scFv fragment for the hapten phenyloxazolone derived from the phage library (and like the anti-self antibodies described here with a relatively fast off-rate) was improved from $K_a = 3.1\times10^6\ M^{-1}$ to $9.1\times10^8\ M^{-1}$ by chain shuffling (Marks *et al.*, 1992b). This would allow the creation of highly specific, high affinity human antibodies directed against self-antigens for use in human therapy.

## Materials and methods

### *Selection of phage library*

The construction and selection of the library of phages displaying scFv fragments has been described previously by Marks *et al.* (1991) and as briefly summarized in the Introduction. The library used in this work contained $2.9\times10^7$ clones and the VH genes were amplified from cDNA primed with an IgM-specific constant region primer.

To rescue the library 50 ml of 2×TY broth (Miller, 1972) containing 100 μg ampicillin/ml, 1% glucose (2×TY-AMP-GLU) were inoculated with $10^9$ *E.coli* TG1; (Gibson, 1984) of the library stock (~10 μl) and grown, shaking at 37°C until the culture reached an $OD_{600\ nm}$ of 0.5. 5 ml of this culture (~$2.5\times10^9$ cells) were added to 50 ml 2×TY-AMP-GLU (pre-warmed to 37°C) containing $5\times10^{10}$ p.f.u. VCS-M13 helper phage (Stratagene) or $5\times10^8$ pfu M13ΔgIII helper phage (unpublished data). M13ΔgIII helper phage does not encode pIII hence the phage(mid) displaying antibody fragments have a greater avidity of binding to antigen. Infectious M13ΔgIII particles were made by growing the helper phage in cells harbouring a pUC19 derivative supplying the wild-type gIII protein during phage morphogenesis. The culture was incubated for 1 h at 37°C without shaking and then a further 1 h at 37°C with shaking. Cells were spun down (IEC-Centra 8, 4000 r.p.m. for 10 min), resuspended in 300 ml 2×TY broth containing 100 μg/ml ampicillin and 25 μg/ml kanamycin (2×TY-AMP-KAN) and grown overnight shaking at 37°C (or 25°C for phage selected with rsCD4). Phage particles were purified and concentrated from the culture medium by two PEG precipitations (Sambrook *et al.*, 1990), resuspended in 2 ml PBS and passed through a 0.45 μM filter (Minisart NML; Sartorius) to give a final concentration of ~$10^{13}$ transducing units/ml (ampicillin-resistant clones).

The phage were panned for binding using immuno tubes (Nunc; Maxisorp) coated with antigen essentially as Marks *et al.* (1991), or were selected on a column of antigen (McCafferty *et al.*, 1990). Six antigens were used: a human mAb Fog-1 (γ1, κ) (Melamed *et al.*, 1987); recombinant human tumour necrosis factor-α (TNFα) (expressed in yeast); bovine thyroglobulin (Sigma); human recombinant soluble CD4 (rsCD4) (expressed in baculovirus by American Biotechnologies Inc. and supplied by the MRC AIDS Reagent Project [ADP608]); human carcinoembryonic antigen (CEA); and a 20 amino acid peptide (Price *et al.*, 1990), which corresponds to a repeated motif in human MUC1 mucin (tumour-associated polymorphic epithelial mucin or PEM) (Gendler *et al.*, 1988; Gum *et al.*, 1990). All antigens, except the MUC1 peptide, were coated on immuno tubes overnight at room temperature. 10 μg/ml TNFα was coated in 50 mM $NaHCO_3$ (pH 9.6), whilst the other antigens were coated in PBS at a concentration of 10 μg/ml Fog1 and rsCD4, 1 mg/ml bovine thyroglobulin or 20 μg/ml CEA. For the first two rounds of selection tubes were washed 10 times with PBS, 0.1% (v/v) Tween 20 and 10 times with PBS. For subsequent rounds of selection tubes were washed 20 times with PBS, 0.1% (v/v) Tween 20 and 20 times with PBS. Phage were eluted with 100 mM triethylamine as Marks *et al.* (1991). Eluted phage (usually $10^6$-$10^7$ transducing units) were used to infect *E.coli* TG1 cells. Approximately $10^9$ infected bacteria were used as an inoculum for the next rescue. The library was subjected to 3–5 rounds of rescue and selection for each antigen.

For selection of phage binding to the MUC1 peptide, the peptide was coupled chemically to Sepharose 4B (provided by M.R.Price). A 1 ml column was prepared and phage was selected as described by McCafferty *et al.* (1990). Briefly, the Sepharose–MUC1 column was washed with PBS containing 2% skimmed milk powder (MPBS) and the phage loaded in 1ml of the same buffer. After washing the column successively with 10 ml volumes of MPBS, PBS (pH 7.2), 50 mM Tris–HCl–500 mM NaCl (pH 8.0) and 50 mM Tris–HCl–500 mM NaCl (pH 9.0), phage was eluted with 5 ml 100 mM triethylamine and neutralized with 0.5 M sodium phosphate buffer (pH 6.8). Five rounds of selection were carried out.

### *Screening and sequencing of clones*

Single ampicillin resistant colonies from infection of *E.coli* TG1 with eluted phage, were screened either for binding of phage (Clackson *et al.*, 1991) or soluble scFv fragments (Marks *et al.*, 1991). Since the gene encoding the antibody fragment is linked to that encoding the phage coat protein by an amber codon, soluble fragments can be secreted from a non-suppressor strain of bacteria infected by the phage (Hoogenboom *et al.*, 1991). The binding to antigen of soluble scFvs in bacterial supernatant was detected with the mouse mAb 9E10 (1 μg/ml), which recognizes the C-terminal peptide tag (Munro and Pelham, 1986) and peroxidase-conjugated anti-mouse Fc antibody (Sigma), as described by Ward *et al.* (1989). Plates were coated with the antigens Fog1, TNFα, bovine thyroglobulin and rsCD4 as described for immuno tubes above and with CEA at 5 mg/ml. A urine extract containing human polymorphic epithelial mucin (PEM) was used at a protein concentration of ~10 mg/ml.

The specificity of the isolated clones was checked by ELISA of the soluble scFv fragments using plates coated with various proteins. Plates were coated with the antigens Fog-1, TNFα, bovine thyroglobulin, rsCD4, CEA and PEM as described above. Other proteins were coated overnight at room temperature at a concentration of 1 mg/ml in PBS (cytochrome *c* [Sigma]) or in 50 mM $NaHCO_3$ (pH 9.6) and bovine serum albumin, turkey egg

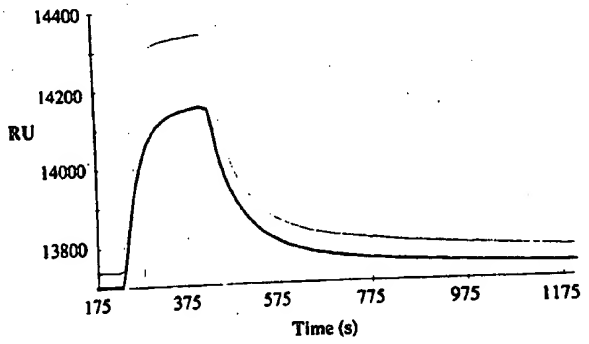



**Fig. 5.** BIAcore sensorgram [resonance signal (RU) as a function of time] of the interaction of soluble αFOG1-H6 scFv dimer with immobilized mAb Fog-1 (Melamed *et al.*, 1987). A 35 μl pulse of 200 nM (thin line) or 80 nM (thick line) scFv protein was passed, with a flow rate of 10 μl/min over a sensor chip to which mAb Fog-1 was coupled directly.

foreign antigens from unimmunized human donors (Marks *et al.*, 1991). However, the origins of the V-genes of scFv fragments directed against self-antigens are less clear. Self-reactive antibodies, including those with specificities against human thyroglobulin (Ruf *et al.*, 1985), human TNFα (Bendtzen *et al.*, 1990) and human IgG (Welch *et al.*, 1983), are common in healthy individuals and indeed 10–30% of B-lymphocytes appear to be engaged in making autoantibodies (Cohen and Cooke, 1986). Therefore the V-genes could be derived from B-cells that are autoreactive, or those that are not. Since somatic hypermutation of antibody genes is triggered only after antigen-induced B-cell proliferation (Griffiths *et al.*, 1984), the isolation of scFv fragments encoded by somatically mutated V-genes (Table III) indicates that the V-genes have been derived from lymphocytes that have been stimulated by antigen; for example from B-cells with self-specificities that have been stimulated with cross-reactive foreign antigen, or from B-cells encoding antibodies of other (foreign) specificities. Conversely those scFv fragments encoded by V-genes with little or no somatic mutation (see Table III) may well have been derived from virgin B-cells or those involved in early immune responses.

'Natural autoantibodies' (self-reactive antibodies isolated from healthy donors) tend to be of low affinity and polyspecific and may well be produced by a discrete subset of B-cells, the internal activity set (Holmberg and Coutinho, 1985), contributed in part by $CD5^+$ B-cells (Casali and Notkins, 1989). In contrast, these anti-self scFv fragments are highly specific in binding to antigen despite only having micromolar affinities. However, their affinities could presumably be improved *in vitro*, for example, the affinity of an scFv fragment for the hapten phenyloxazolone derived from the phage library (and like the anti-self antibodies described here with a relatively fast off-rate) was improved from $K_a = 3.1\times10^6\ M^{-1}$ to $9.1\times10^8\ M^{-1}$ by chain shuffling (Marks *et al.*, 1992b). This would allow the creation of highly specific, high affinity human antibodies directed against self-antigens for use in human therapy.

## Materials and methods

### *Selection of phage library*

The construction and selection of the library of phages displaying scFv fragments has been described previously by Marks *et al.* (1991) and as briefly summarized in the Introduction. The library used in this work contained $2.9\times10^7$ clones and the VH genes were amplified from cDNA primed with an IgM-specific constant region primer.

To rescue the library 50 ml of 2×TY broth (Miller, 1972) containing 100 μg ampicillin/ml, 1% glucose (2×TY-AMP-GLU) were inoculated with $10^9$ *E.coli* TG1; (Gibson, 1984) of the library stock (~10 μl) and grown, shaking at 37°C until the culture reached an $OD_{600\ nm}$ of 0.5. 5 ml of this culture (~$2.5\times10^9$ cells) were added to 50 ml 2×TY-AMP-GLU (pre-warmed to 37°C) containing $5\times10^{10}$ p.f.u. VCS-M13 helper phage (Stratagene) or $5\times10^8$ pfu M13ΔgIII helper phage (unpublished data). M13ΔgIII helper phage does not encode pIII hence the phage(mid) displaying antibody fragments have a greater avidity of binding to antigen. Infectious M13ΔgIII particles were made by growing the helper phage in cells harbouring a pUC19 derivative supplying the wild-type gIII protein during phage morphogenesis. The culture was incubated for 1 h at 37°C without shaking and then a further 1 h at 37°C with shaking. Cells were spun down (IEC-Centra 8, 4000 r.p.m. for 10 min), resuspended in 300 ml 2×TY broth containing 100 μg/ml ampicillin and 25 μg/ml kanamycin (2×TY-AMP-KAN) and grown overnight shaking at 37°C (or 25°C for phage selected with rsCD4). Phage particles were purified and concentrated from the culture medium by two PEG precipitations (Sambrook *et al.*, 1990), resuspended in 2 ml PBS and passed through a 0.45 μM filter (Minisart NML; Sartorius) to give a final concentration of ~$10^{13}$ transducing units/ml (ampicillin-resistant clones).

The phage were panned for binding using immuno tubes (Nunc; Maxisorp) coated with antigen essentially as Marks *et al.* (1991), or were selected on a column of antigen (McCafferty *et al.*, 1990). Six antigens were used: a human mAb Fog-1 (γ1, κ) (Melamed *et al.*, 1987); recombinant human tumour necrosis factor-α (TNFα) (expressed in yeast); bovine thyroglobulin (Sigma); human recombinant soluble CD4 (rsCD4) (expressed in baculovirus by American Biotechnologies Inc. and supplied by the MRC AIDS Reagent Project [ADP608]); human carcinoembryonic antigen (CEA); and a 20 amino acid peptide (Price *et al.*, 1990), which corresponds to a repeated motif in human MUC1 mucin (tumour-associated polymorphic epithelial mucin or PEM) (Gendler *et al.*, 1988; Gum *et al.*, 1990). All antigens, except the MUC1 peptide, were coated on immuno tubes overnight at room temperature. 10 μg/ml TNFα was coated in 50 mM $NaHCO_3$ (pH 9.6), whilst the other antigens were coated in PBS at a concentration of 10 μg/ml Fog1 and rsCD4, 1 mg/ml bovine thyroglobulin or 20 μg/ml CEA. For the first two rounds of selection tubes were washed 10 times with PBS, 0.1% (v/v) Tween 20 and 10 times with PBS. For subsequent rounds of selection tubes were washed 20 times with PBS, 0.1% (v/v) Tween 20 and 20 times with PBS. Phage were eluted with 100 mM triethylamine as Marks *et al.* (1991). Eluted phage (usually $10^6$-$10^7$ transducing units) were used to infect *E.coli* TG1 cells. Approximately $10^9$ infected bacteria were used as an inoculum for the next rescue. The library was subjected to 3–5 rounds of rescue and selection for each antigen.

For selection of phage binding to the MUC1 peptide, the peptide was coupled chemically to Sepharose 4B (provided by M.R.Price). A 1 ml column was prepared and phage was selected as described by McCafferty *et al.* (1990). Briefly, the Sepharose–MUC1 column was washed with PBS containing 2% skimmed milk powder (MPBS) and the phage loaded in 1ml of the same buffer. After washing the column successively with 10 ml volumes of MPBS, PBS (pH 7.2), 50 mM Tris–HCl–500 mM NaCl (pH 8.0) and 50 mM Tris–HCl–500 mM NaCl (pH 9.0), phage was eluted with 5 ml 100 mM triethylamine and neutralized with 0.5 M sodium phosphate buffer (pH 6.8). Five rounds of selection were carried out.

### *Screening and sequencing of clones*

Single ampicillin resistant colonies from infection of *E.coli* TG1 with eluted phage, were screened either for binding of phage (Clackson *et al.*, 1991) or soluble scFv fragments (Marks *et al.*, 1991). Since the gene encoding the antibody fragment is linked to that encoding the phage coat protein by an amber codon, soluble fragments can be secreted from a non-suppressor strain of bacteria infected by the phage (Hoogenboom *et al.*, 1991). The binding to antigen of soluble scFvs in bacterial supernatant was detected with the mouse mAb 9E10 (1 μg/ml), which recognizes the C-terminal peptide tag (Munro and Pelham, 1986) and peroxidase-conjugated anti-mouse Fc antibody (Sigma), as described by Ward *et al.* (1989). Plates were coated with the antigens Fog1, TNFα, bovine thyroglobulin and rsCD4 as described for immuno tubes above and with CEA at 5 mg/ml. A urine extract containing human polymorphic epithelial mucin (PEM) was used at a protein concentration of ~10 mg/ml.

The specificity of the isolated clones was checked by ELISA of the soluble scFv fragments using plates coated with various proteins. Plates were coated with the antigens Fog-1, TNFα, bovine thyroglobulin, rsCD4, CEA and PEM as described above. Other proteins were coated overnight at room temperature at a concentration of 1 mg/ml in PBS (cytochrome *c* [Sigma]) or in 50 mM $NaHCO_3$ (pH 9.6) and bovine serum albumin, turkey egg-

white lysozyme, hen egg-white lysozyme, hen ovalbumin, keyhole limpet haemocyanin (CalBiochem), chymotrypsinogen A, chicken egg-white trypsin inhibitor (Sigma) and chicken γ globulin coupled to 4-hydroxy-3-nitrophenyl acetic acid. The Fog-1 specific clones were screened by binding to a panel of different human antibodies (see legend to Figure 3). The antibodies were coated overnight at room temperature in PBS at a concentration of 10 μg/ml.

Clones found to give a positive ELISA signal were screened by PCR (Gussow and Clackson, 1989) and 'fingerprinted' with the restriction enzyme *Bst*NI (Clackson *et al.*, 1991) as in Marks *et al.* (1991) to identify different clones. Examples of clones with different restriction patterns were selected and the heavy and light chains sequenced (Sanger *et al.*, 1977) using a Sequenase kit (USB) or using a Taq DyeDeoxy Terminator Cycle Sequencing kit (Applied Biosystems) and an Applied Biosystems 373A DNA sequencer.

Sequenced clones were further analysed using the program MacVector 3.5 (IBI Kodak, New Haven, CT). The VH genes were compared with the 83 germline gene segments present in the VH directory compiled by Tomlinson *et al.* (1992). VL genes were compared with 34 published kappa germline gene segments (Bentley and Rabbits, 1980, 1983; Jaenichen *et al.*, 1984; Pech *et al.*, 1984, 1985; Pech and Zachau, 1984; Klobeck *et al.*, 1985a,b; Stavnezer *et al.*, 1985; Chen *et al.*, 1986, 1987a,b; Lorenz *et al.*, 1988; Straubinger *et al.*, 1988a,b; Scott *et al.*, 1989, 1991; Pargent *et al.*, 1991; Lautner *et al.*, 1992) and 13 published lambda gene segments (Anderson *et al.*, 1984; Alexandre *et al.*, 1989; Brockly *et al.*, 1989; Siminovitch *et al.*, 1989; Bernard *et al.*, 1990; Frippiat *et al.*, 1990; Combriato and Klobeck, 1991; Marks *et al.*, 1991; Daley *et al.*, 1992; Winkler *et al.*, 1992). Regions of the V-genes encoded by PCR primers were not included in the analysis.

### *Characterization of selected scFv fragments*

The following clones were chosen for large scale purification and further characterization: αFOG1-H6, αFOG1-A3, αTNF-E7 and αThy-29. Colonies of the non-suppressor *E.coli* strain HB2151 harbouring the appropriate phagemid were used to inoculate 2 l of 2×TY containing 100 μg/ml ampicillin and 0.1% glucose. The cultures were grown and induced (De Bellis and Schwartz, 1990) and the tagged scFv fragments purified using the mAb 9E10 as in Clackson *et al.* (1991).

The inhibition of $^{125}$I-Fog-1 binding to human Rh D antigen by the affinity purified scFv fragments αFOG1-H6 and αFOG1-A3 was essentially as performed earlier (Gorick *et al.*, 1988) with the following modifications. 0.0148 μg of $^{125}$I-FOG1 was pre-incubated with varying amounts of purified αFOG1-H6 or αFOG1-A3 scFv fragments (0–16 μg) at 37°C for 1.5 h, before adding 0.5 μl of $R_1R_2$ cells (or rr cells as control). The mixture was then incubated for a further 1.5 h at 37°C with constant mixing and finally cells separated from the supernatant. As a control, a titration was also performed with a purified scFv fragment directed against turkey egg white lysozyme (αTEL9) (Marks *et al.*, 1991).

Kinetic measurements were made using surface plasmon resonance (BIAcore, Pharmacia Biosensor AB) (Jönsson *et al.*, 1991; Jönsson and Malmqvist, 1992). In order to separate monomeric and multimeric species, the purified scFv fragments were concentrated by ultrafiltration and then fractionated on a calibrated Superdex 75 FPLC column (Pharmacia) in PBS, 0.2 mM EDTA. Gel filtration was monitored both by the absorbance at 280 nm and on-line to BIAcore with immobilized antigen on the sensor chip (Johnsson *et al.*, 1991).

Kinetic experiments were performed in two different configurations. First, to analyse the binding of soluble scFv, the different antigens were covalently immobilized on the sensor chip (in the case of mAb Fog-1, the antibody was also immobilized via a mouse anti-human kappa light chain mAb using a sensor chip coated with rabbit anti-mouse IgG1). Secondly, to analyse the binding of the soluble mAb FOG-1, the αFOG1-H6 scFv was immobilized on the chip surface.

The antigens were coupled to the CM5 sensor chip through their amine groups using the Amine Coupling Kit (Pharmacia Biosensor AB) (Johnsson *et al.*, 1991). The antigens were diluted in 10 mM acetate buffer (pH 5.0) to ~25 μg/ml and 3805 resonance units (RU) of TNF, 6249 RU of human thyroglobulin and 5279 RU of FOG1 were immobilized. For the biospecific presentation of Fog-1, affinity purified rabbit anti-mouse IgG1 (Pharmacia Biosensor AB) was coupled to the surface followed by a mouse mAb anti-human kappa (2300 RU) and then Fog-1 (2050 RU). As binding of the rabbit anti-mouse IgG1 to the mouse mAb was reversible by 10 mM HCl the complex was rebuilt for each analytical cycle. ScFv anti-Fog-1 was coupled to the CM5 surface to 1538 RU. All determinations were performed at 25°C in PBS, 0.2 mM EDTA, 0.05% BIAcore surfactant P20 with a constant flow rate of 10 μl/min and an injected vol sample of 35 μl. It was not necessary to regenerate the antigen as the scFv fragments rapidly dissociate, with the exception of the biospecific presentation of antigen via rabbit anti-mouse IgG1 which was regenerated with 10 mM HCl for 3 min.

Analyses of scFv monomer were performed in the concentration range 100–500 nM and dimers in the range 40–200 nM except for the biospecifically presented Fog-1 where the concentration of dimeric scFv was 0.25–1.26 μM. Fog-1 was analysed on the αFOG1-H6 scFv surface in the concentration range 10–200 nM. All concentrations were calculated from U.V. absorption at 280 nm [assuming that 0.7 mg/ml scFv gives an $A_{280}$ = 1 (Mach *et al.*, 1992) and that $M_r$ of a scFv monomer is 30 kDa and of a dimer is 60 kDa]. No correction was made for the fraction of active protein, and therefore the on-rates are an underestimate. The kinetic evaluation of data was performed according to Karlsson *et al.* (1991) and evaluated on the program Origin 1.1 (Microcal inc., Northampton, MA, USA).

## Acknowledgements

We thank M.R.Price for providing CEA, MUC1 peptide coupled to Sepharose 4B and PEM, A.Nissim, C.Chothia and S.Williams for help with sequence analysis, and Peptide Technology Ltd (Sydney, Australia) for a gift of recombinant TNFα. Recombinant soluble CD4 was provided by the MRC AIDS Reagent Project. A.D.G. and M.J.E. were supported by the Cancer Research Campaign, M.M. by Pharmacia Biosensor AB, J.D.M. by the MRC AIDS Directed Programme and the MRC, J.M.B., B.D.G. and N.C.H.-J. by a grant from the International Blood Group Reference Laboratory, J.McC. by Cambridge Antibody Technology Ltd., M.B. by the AIDS Program of the Federal Ministry for Research and Technology Germany, K.P.H. by ETH Zurich, and H.R.H. by the D.Collen Research Foundation, Leuven and the European Molecular Biology Organization.

## References

Adelstein,S., Pritchard,B.H., Anderson,T.A., Crosbie,J., Gammon,G., Loblay,R.H., Basten,A. and Goodnow,C.C. (1991) *Science*, **251**, 1223–1225.

Alexandre,D., Chuchana,P., Brockly,F., Blancher,A., Lefranc,G. and Lefranc,M.P. (1989) *Nucleic Acids Res.*, **17**, 3975.

Anderson,M.L., Szajnert,M.F., Kaplan,J.C., McColl,L. and Young,B.D. (1984) *Nucleic Acids Res.*, **12**, 6647–6661.

Avrameas,S. (1991) *Immunol. Today*, **12**, 154–159.

Basten,A., Brink,R., Peake,P., Adams,E., Crosbie,J., Hartley,S. and Goodnow,C.C. (1991) *Immunol. Rev.*, **122**, 5–19.

Bendtzen,K., Svenson,M., Jnsson,V. and Hippe,E. (1990) *Immunol. Today*, **11**, 167–169.

Bentley,D.L. and Rabbits,T.H. (1980) *Nature*, **288**, 730–733.

Bentley,D.L. and Rabbitts,T.H. (1983) *Cell*, **32**, 181–189.

Berek,C. and Milstein,C. (1987) *Immunol. Rev.*, **96**, 23–41.

Bernard,F., Chuchana,P., Frippiat,J.P., Buluwela,L. and Lefranc,M.P. (1990) *Nucleic Acids Res.*, **18**, 7139.

Bird,R.E. *et al.* (1988) *Science*, **242**, 423–426.

Bona,C.A. (1988) *Annu. Rev. Immunol.*, **6**, 327–358.

Bouanani,M., Bataille,R., Piechaczyk,M., Salhi,S.L., Pau,B. and Bastide,M. (1991) *Arthritis Rheumatol.*, **34**, 1585–1593.

Bretscher,P. and Cohn,M. (1970) *Science*, **169**, 1042–1049.

Brockly,F., Alexandre,D., Chuchana,P., Huck,S., Lefranc,G. and Lefranc,M.P. (1989) *Nucleic Acids Res.*, **17**, 3976.

Burnet,F.M. (1959) *The Clonal Selection Theory of Acquired Immunity*. Vanderbilt University Press, Nashville.

Bye,J.M., Carter,C., Cui,Y., Gorick,B.D., Songsivilai,S., Winter,G., Hughes-Jones,N.C. and Marks,J.D. (1993) *J. Clin. Invest.*, in press.

Casali,P. and Notkins,A.L. (1989) *Annu. Rev. Immunol.*, **7**, 513–535.

Chen,P.P., Albrandt,K., Orida,N.K., Radoux,V., Chen,E.Y., Schrantz,R., Liu,F.T. and Carson,D.A. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 8318–8322.

Chen,P.P., Albrandt,K., Kipps,T.J., Radoux,V., Liu,F.-T. and Carson,D.A. (1987a) *J. Immunol.*, **139**, 1727–1733.

Chen,P.P., Robbins,D.L., Jirik,F.R., Kipps,T.J. and Carson,D.A. (1987b) *J. Exp. Med.*, **166**, 1900–1905.

Clackson,T., Hoogenboom,H.R., Griffiths,A.D. and Winter,G. (1991) *Nature*, **352**, 624–628.

Cohen,I.R. and Cooke,A. (1986) *Immunol. Today*, **7**, 363–364.

Combriato,G. and Klobeck,H.G. (1991) *Eur. J. Immunol.*, **21**, 1513–1522.

Daley,D.M., Olee,T., Peng,H.-Q., Soto-Gil,R.W., Chen,P.P. and Siminovitch,K.A. (1992) *Mol. Immunol.*, **29**, 1031–1042.

De Bellis,D. and Schwartz,I. (1990) *Nucleic Acids Res.*, **18**, 1311.

Erikson,J., Radic,M.Z., Camper,S.A., Hardy,R.R., Carmack,C. and Weigert,M. (1991) *Nature*, **349**, 331–334.

Foote,J. and Milstein,C. (1991) *Nature*, **352**, 530–532.

Foote,J. and Winter,G. (1992) *J. Mol. Biol.*, **224**, 487–499.

Frippiat,J.P., Chuchana,P., Bernard,F., Buluwela,L., Lefranc,G. and Lefranc,M.P. (1990) *Nucleic Acids Res.*, 18, 7134.
Gendler,S., Taylor-Papadimitrou,J., Duhig,T., Rothbard,J. and Burchell,J. (1988) *J. Biol. Chem.*, 263, 12820–12823.
Gherardi,E. and Milstein,C. (1992) *Nature*, 357, 201–202.
Gibson,T.J. (1984) PhD thesis, University of Cambridge, UK.
Glockshuber,R., Malia,M., Pfitzinger,I. and Pluckthun,A. (1990) *Biochemistry*, 29, 1362–1367.
Gorick,B.D., Thompson,K.M., Melamed,M.D. and Hughes,J.N. (1988) *Vox. Sang.*, 55, 165–170.
Gram,H., Marconi,L., Barbas,C.F., Collet,T.A., Lerner,R.A. and Kang,A.S. (1992) *Proc. Natl. Acad. Sci. USA*, 89, 3576–3580.
Griffiths,G.M., Berek,C., Kaartinen,M. and Milstein,C. (1984) *Nature*, 312, 271–275.
Gum,J.R., Hicks,J.W., Swallow,D.M., Lagace,R.L., Byrd,J.C., Lamport,D.T.A., Siddiki,B. and Kim,Y.S. (1990) *Biochem. Biophys Res. Commun.*, 171, 407–415.
Gussow,D. and Clackson,T. (1989) *Nucleic Acids Res.*, 17, 4000.
Hale,G., Dyer,M.J., Clark,M.R., Phillips,J.M., Marcus,R., Riechmann,L., Winter,G. and Waldmann,H. (1988) *Lancet*, 2, 1394–1399.
Hartley,S.B., Crosbie,J., Brink,R., Kantor,A.B., Basten,A. and Goodnow,C.C. (1991) *Nature*, 353, 765–769.
Hawkins,R.E. and Winter,G. (1992) *Eur. J. Immunol.*, 22, 867–870.
Holmberg,D. and Coutinho,A. (1985) *Immunol. Today*, 6, 356–357.
Hoogenboom,H.R., Griffiths,A.D., Johnson,K.S., Chiswell,D.J., Hudson,P. and Winter,G. (1991) *Nucleic Acids Res.*, 19, 4133–4137.
Hoogenboom,H.R., Marks,J.D., Griffiths,A.D. and Winter,G. (1992) *Immunol. Rev.*, in press.
Huse,W.D., Sastry,L., Iverson,S.A., Kang,A.S., Alting,M.M., Burton,D.R., Benkovic,S.J. and Lerner,R.A. (1989) *Science*, 246, 1275–1281.
Huston,J.S. *et al.* (1988) *Proc. Natl. Acad. Sci. USA*, 85, 5879–5883.
Jaenichen,H.R., Pech,M., Lindenmaier,W., Wildgruber,N. and Zachau,H.G. (1984) *Nucleic Acids Res.*, 12, 5249–5263.
James,K. and Bell,G.T. (1987) *J. Immunol. Methods*, 100, 5–40.
Johnsson,B., Löfås,S. and Lindqvist,G. (1991) *Anal. Biochem.*, 198, 268–277.
Jones,D.H., McMillan,A.J., Fersht,A.R. and Winter,G. (1985) *Biochemistry*, 24, 5852–5857.
Jönsson,U. and Malmqvist,M. (1992) In Turner,A. (ed.), *Real Time Biospecific Interaction*. JAI Press Ltd, San Diego, Vol 2, pp. 291–336.
Jönsson,U. *et al.* (1991) *BioTechniques*, 11, 620–627.
Karlsson,R., Michaelsson,A. and Mattsson,L. (1991) *J. Immunol. Methods*, 145, 229–240.
Klobeck,H.G., Bornkamm,G.W., Combriato,G., Mocikat,R., Pohlenz,H.D. and Zachau,H.G. (1985a) *Nucleic Acids Res.*, 13, 6515–6529.
Klobeck,H.G., Meindl,A., Combriato,G., Solomon,A. and Zachau,H.G. (1985b) *Nucleic Acids Res.*, 13, 6499–6513.
Laemmli,U.K. (1970) *Nature*, 227, 680–685.
Lautner,R.A., Huber,C., Meindl,A., Pargent,W., Schable,K.F., Thiebe,R., Zocher,I. and Zachau,H.G. (1992) *Eur. J. Immunol.*, 22, 1023–1029.
Lorenz,W., Schable,K.F., Thiebe,R., Stavnezer,J. and Zachau,H.G. (1988) *Mol. Immunol.*, 25, 479–484.
Mach,H., Middaugh,C.R. and Lewis,R.V. (1992) *Anal. Biochem.*, 200, 74–80.
Mach,J.-P., Forni,M., Ritschard,J., Buchegger,F., Carrel,S., Widgren,S., Donath,A. and Alberto,P. (1980) *Oncodevel. Biol. Med.*, 1, 49–69.
Malthiéry,Y. and Lissitzky,S. (1987) *Eur. J. Biochem.*, 165, 491–498.
Marks,J.D., Hoogenboom,H.R., Bonnert,T.P., McCafferty,J., Griffiths,A.D. and Winter,G. (1991) *J. Mol. Biol.*, 222, 581–597.
Marks,J.D., Hoogenboom,H.R., Griffiths,A.D. and Winter,G. (1992a) *J. Biol. Chem.* 267, 16007–16010.
Marks,J.D., Griffiths,A.D., Malmqvist,M., Clackson,T., Bye,J.M. and Winter,G. (1992b) *Bio/Technology*, 10, 779–783.
Mason,D.W. and Williams,A.F. (1986) *Kinetics of Antibody Reactions and the Analysis of Cell Surface Antigens*. Blackwell Scientific, Oxford.
McCafferty,J., Griffiths,A.D., Winter,G. and Chiswell,D.J. (1990) *Nature*, 348, 552–554.
Melamed,M.D., Thompson,K.M., Gibson,T. and Hughes-Jones,N.C. (1987) *J. Immunol. Methods*, 104, 245–251.
Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Munro,S. and Pelham,H.R.B. (1986) *Cell*, 46, 291–300.
Nakamura,M., Burastero,S.E., Ueki,Y., Larrick,J.W., Notkins,A.L. and Casali,P. (1988) *J. Immunol.*, 141, 4165–4172.
Nemazee,D., Russell,D., Arnold,B., Haemmerling,G., Allison,J., Miller,J.F., Morahan,G. and Buerki,K. (1991) *Immunol. Rev.*, 122, 117–132.
Nossal,G.J. (1987) *Int. Rev. Immunol.*, 2, 321–338.
Nossal,G.J. (1989) *Science*, 245, 147–153.
Pargent,W., Meindl,A., Thiebe,R., Mitzel,S. and Zachau,H.G. (1991) *Eur. J. Immunol.*, 21, 1821–1827.
Pech,M. and Zachau,H.G. (1984) *Nucleic Acids Res.*, 12, 9229–9236.
Pech,M., Jaenichen,H.R., Pohlenz,H.D., Neumaier,P.S., Klobeck,H.G. and Zachau,H.G. (1984) *J. Mol. Biol.*, 176, 189–204.
Pech,M., Smola,H., Pohlenz,H.D., Straubinger,B., Gerl,R. and Zachau,H.G. (1985) *J. Mol. Biol.*, 183, 291–299.
Pecht,I. (1982) In Sela,M. (ed.), *Dynamic Aspects of Antibody Function*. Academic Press, Inc., New York, Vol 6, pp. 1–68.
Perelson,A.S. and Oster,G.F. (1979) *J. Theor. Biol.*, 81, 645–670.
Persson,M.A.A., Caothien,R.H. and Burton,D.R. (1991) *Proc. Natl. Acad. Sci. USA*, 88, 2432–2436.
Price,M.R., Hudecz,F., O'Sullivan,C., Baldwin,R.W., Edwards,P.M. and Tendler,S.J.B. (1990) *Mol. Immunol.*, 27, 795–802.
Rossi,F., Guilbert,B., Tonnelle,C., Ternynck,T., Fumoux,F., Avrameas,S. and Kazatchkine,M.D. (1990) *Eur. J. Immunol.*, 20, 2089–2094.
Ruf,J., Carayon,P. and Lissitzky,S. (1985) *Eur. J. Immunol.*, 15, 268–272.
Russell,D.M., Dembic,Z., Morahan,G., Miller,J.F., Burki,K. and Nemazee,D. (1991) *Nature*, 354, 308–311.
Sambrook,J., Fritsch,E.F. and Maniatis,T. (1990) *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, 74, 5463–5467.
Scott,M.G., Crimmins,D.L., McCourt,D.W., Zocher,I., Thiebe,R., Zachau,H.G. and Nahm,M.H. (1989) *J. Immunol.*, 143, 4110–4116.
Scott,M.G., Crimmins,D.L., McCourt,D.W., Chung,G., Schable,K.F., Thiebe,R., Quenzel,E.M., Zachau,H.G. and Nahm,M.H. (1991) *J. Immunol.*, 147, 4007–4013.
Siminovitch,K.A., Misener,V., Kwong,P.C., Song,Q.-L. and Chen,P.P. (1989) *J. Clin. Invest.*, 84, 1675–1678.
Smith,G.P. (1985) *Science*, 228, 1315–1317.
Smith,H.R. and Steinberg,A.D. (1983) *Annu. Rev. Immunol.*, 1, 175–210.
Smith,T.W. and Skubitz,K.M. (1975) *Biochemistry*, 14, 1496–1502.
Spooner,C.E., Markowitz,N.P. and Saravolatz,L.D. (1992) *Clin. Immunol. Immunopathol.*, 62, S11–17.
Stavnezer,J., Kekish,O., Batter,D., Grenier,J., Balazs,I., Henderson,E. and Zegers,B.J. (1985) *Nucleic Acids Res.*, 13, 3495–514.
Straubinger,B., Huber,E., Lorenz,W., Osterholzer,E., Pargent,W., Pech,M., Pohlenz,H.D., Zimmer,F.J. and Zachau,H.G. (1988a) *J. Mol. Biol.*, 199, 23–34.
Straubinger,B., Thiebe,R., Huber,C., Osterholzer,E. and Zachau,H.G. (1988b) *Biol. Chem. Hoppe-Seyler*, 369, 601–607.
Thompson,K.M., Hough,D.W., Maddison,P.J., Melamed,M.D. and Hughes-Jones, N.C. (1986) *J. Immunol. Methods*, 94, 7–12.
Tomer,Y. and Schoenfeld,Y. (1988) *Immunol. Invest.*, 17, 389–424.
Tomlinson,I.M., Walter,G., Marks,J.D., Llewelyn,M.B. and Winter,G. (1992) *J. Mol. Biol.*, 227, 776–798.
Ward,E.S., Gussow,D., Griffiths,A.D., Jones,P.T. and Welch,M.J., Fong,S., Vaughan,J. and Carson,D. (1983) *Clin. Exp. Immunol.*, 51, 299–304.
Winkler,T.H., Fehr,H. and Kalden,J.R. (1992) *Eur. J. Immunol.*, 22, 1719–1728.
Winter,G. (1989) *Nature*, 341, 544–546.
Winter,G. and Milstein,C. (1991) *Nature*, 349, 293–299.

*Received on September 29, 1992; revised on November 11, 1992*

## Note added in proof

Nucleotide sequences of all the V-genes whose deduced protein sequences are given in Table II have been submitted to the EMBL Data Library and assigned the following accession numbers: FOG1VHA3, Z18822; FOG1VLA3, Z18823; FOG1VHH6, Z18824; FOG1VLH6, Z18825; FOG1VHA4, Z18826; FOG1VLA4, Z18827; FOG1VHG8, Z18828; FOG1VLG8, Z18829; THYVH23, Z18830; THYVL23, Z18831; THYVH29, Z18832; THYVL29, Z18833; THYVH32, Z18834; THYVL32, Z18835; THYVH33, Z18836; THYVL33, Z18837; TNFVLA1, Z18838; TNFVHE1, Z18839; TNFVLE1, Z18840; TNFVHE7, Z18841; TNFVLE7, Z18842; TNFVHH9, Z18843; TNFVLH9, Z18844; CEAVL8A, Z18845; MUC1VH1, Z18846; MUC1VL1, Z18847; CD4VH74, Z18848; CD4VL74, Z18849; TNFVHA1, Z18850; CEAVH8A, Z18851.

Protein Engineering vol.10 no.4 pp.435–444, 1997

# Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved *in vivo* folding and physical characterization of an engineered scFv fragment

Lars Nieba, Annemarie Honegger, Claus Krebber and Andreas Plückthun[1]

Biochemisches Institut der Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

[1]To whom correspondence should be addressed

**By constructing Fv and single-chain Fv (scFv) fragments of antibodies, the variable domains are taken out of their natural context in the Fab fragment, where they are associated with the constant domains of the light ($C_L$) and heavy chain ($C_H1$). As a consequence, all residues of the former variable/constant domain interface become solvent exposed. In an analysis of 30 non-redundant Fab structures it was found that at the former variable/constant domain interface of the Fv fragment the frequency of exposed hydrophobic residues is much higher than in the rest of the Fv fragment surface. We investigated the importance of these residues for different properties such as folding *in vivo* and *in vitro*, thermodynamic stability, solubility of the native protein and antigen affinity. The experimental model system was the scFv fragment of the anti-fluorescein antibody 4-4-20, of which only 2% is native when expressed in the periplasm of *Escherichia coli*. To improve its *in vivo* folding, a mutagenesis study of three newly exposed interfacial residues in various combinations was carried out. The replacement of one of the residues (V84D in $V_H$) led to a 25-fold increase of the functional periplasmic expression yield of the scFv fragment of the antibody 4-4-20. With the purified scFv fragment it was shown that the thermodynamic stability and the antigen binding constant are not influenced by these mutations, but the rate of the thermally induced aggregation reaction is decreased. Only a minor effect on the solubility of the native protein was observed, demonstrating that the mutations prevent aggregation during folding and not of the native protein. Since the construction of all scFv fragments leads to the exposure of these residues at the former variable/constant domain interface, this strategy should be generally applicable for improving the *in vivo* folding of scFv fragments and, by analogy, also the *in vivo* folding of other engineered protein domains.**

*Keywords*: antibody engineering/domain interface/hydrophobic patches/protein aggregation/protein engineering/single-chain Fv fragment

## Introduction

Small antibody fragments show exciting promise for use as therapeutics, diagnostic reagents and for biochemical research (Winter and Milstein, 1991; Plückthun, 1994; Huston *et al.*, 1995). For most of these applications they have to be prepared in large amounts. The functional expression of antibody fragments in the periplasm of *Escherichia coli* (Plückthun, 1992), especially Fv or single-chain Fv (scFv) fragments, is now used in many laboratories. Functional expression yields of these fragments vary widely, however, over several orders of magnitude even when the cell density is accounted for or when fragments in the identical host–vector system are compared (Carter *et al.*, 1992; Plückthun *et al.*, 1996). Despite numerous studies (reviewed by Plückthun *et al.*, 1996), the factors influencing antibody expression levels are still only poorly understood. Initial efforts had focused on transcription and translation efficiency, but with present vector systems these problems have been satisfactorily solved. The differences in expression yield between different antibody sequences show that protein sequence-related factors are the major remaining challenge. Folding efficiency and stability of the antibody fragments, as well as protease lability and toxicity of the expressed antibody fragments to the host cells, often severely limit actual production levels, but the rational understanding of these problems is only beginning.

Knappik and Plückthun (1995) showed that the primary sequence of a particular antibody emerges as the most decisive factor determining the yield of functional protein. Based on a sequence comparison of different antibodies, beck-engineering pin-pointed the effect to just a few point mutations. While these particular mutations are only applicable to a few antibodies and thus constitute a tailor-made solution, they do show that very minor sequence changes can have a dramatic effect on the *in vitro* aggregation properties of these molecules and the *in vivo* functional expression of antibody fragments in the periplasm of *E.coli*. Similarly, Ullrich *et al.* (1995) found that point mutations in the complementarity-determining regions (CDRs) can increase the yields in periplasmic antibody fragment expression.

The observations of Knappik and Plückthun (1995) indicate that optimizing those parts of the antibody fragment which are not directly involved in antigen recognition can significantly improve folding properties and production yields of recombinant Fv and scFv constructs. The causes of the improved expression behavior lie in the decreased aggregation behavior of these molecules. The understanding of how specific sequence modifications change these properties is still very limited and currently under active investigation.

Normally, the majority of the side chain residues exposed at the surface of a protein are hydrophilic. Extended hydrophobic patches on the surface often indicate protein–protein interaction sites. If individual protein domains are taken out of their natural biological context, surfaces which are normally buried can become exposed. In the case of the antigen-binding Fv domain of an antibody, the Fv domain is taken away from the constant domains $C_L$ and $C_H1$. While the resulting loss of $V_H/V_L$ interaction energy can be overcome by engineering scFv or disulfide-linked Fv (dsFv) fragments (Huston *et al.*, 1988; Bird *et al.*, 1988; Glockshuber *et al.*, 1990; Brinkmann *et al.*, 1993; Huston *et al.*, 1995; Young *et al.*, 1995), the residues at the former variable/constant (v/c) domain interface, which are usually buried in this interface, remain exposed to

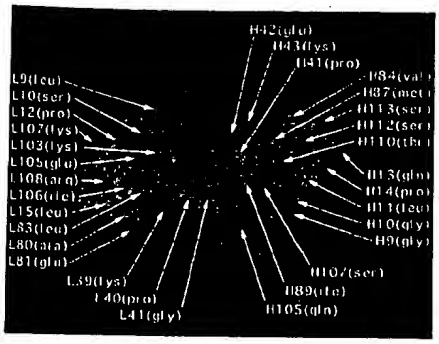



Fig. 1. Space-filling representation of the Fv fragment, seen from the v/c interface, of the antibody 4-4-20, color coded for residue types. Orange, aromatic side chains (Tyr, Phe, Trp); yellow, aliphatic side chains (Leu, Ile, Val, Pro, Ala); sulfur containing side-chains (Met, Cys); green, uncharged, hydrophilic side chains (Thr, Ser, Asn, Gln); red, acidic side chains (Glu, Asp); blue, basic side chains (His, Arg, Lys); white, main-chain (hydrophobicity color code would be yellow–green)

the solvent and form an extended hydrophobic patch (Figure 1). These exposed hydrophobic residues may influence the *in vivo* folding pathway by stabilizing misfolded structures and promoting the aggregation of folding intermediates.

In this study, we investigated a series of mutations introduced into an scFv fragment of the anti-fluorescein antibody 4-4-20 (Bedzyk *et al.*, 1990). In its original form, this fragment is very poorly produced in functional form. Nearly all of the expression product harvested from the periplasm of *E.coli* consists of insoluble, non-functional material (Bedzyk *et al.*, 1990; Denzin *et al.*, 1991). Based on the X-ray stucture of the Fab fragment of this antibody (Whitlow *et al.*, 1995), we identified the hydrophobic residues at the newly exposed v/c domain interface. Substitution of one particular newly exposed hydrophobic residue in this area increases the functional product 25-fold and we characterized the effect on the physical properties of this protein.

## Material and methods

### *Calculation of solvent accessibility*

Solvent-accessible surface areas for 30 non-redundant Fab fragments and the Fv fragments derived from these by deleting the constant domain coordinates from the PDB file were calculated using the latest version of the program NACCESS (http: // www.biochem.ucl.ac.uk / ~roman / naccess / naccess) based on the algorithm described by Lee and Richards (1971).

### *scFv gene synthesis*

The scFv fragment in the orientation $V_L$–linker–$V_H$ of the antibody 4-4-20 (Bedzyk *et al.*, 1990) was obtained by gene synthesis (Prodromou and Pearl, 1992). The $V_L$ domain carries a three amino acid long FLAG tag (Knappik and Plückthun, 1994) and the $His_5$-tag at the end of $V_H$ (Ge *et al.*, 1995). We have used two different linkers of length 15 $(Gly_4Ser)_3$ and 30 amino acids $(Gly_4Ser)_6$, respectively. The gene so obtained was cloned into a derivative of the vector pIG6 (Ge *et al.*, 1995). The mutant antibody fragments were constructed by site-directed mutagenesis (Kunkel *et al.*, 1987) using single-stranded DNA and up to three oligonucleotides per reaction.

### *Expression for measuring in vivo solubility*

Growth curves were obtained as follows: 20 ml of 2×YT medium containing 100 μg/ml ampicillin and 25 μg/ml streptomycin were inoculated with 250 μl of an overnight culture of *E.coli* JM83 harboring the plasmid encoding the respective antibody fragment and incubated at 24.5°C until an $OD_{550}$ of 0.5 was reached. Isopropyl-β-D-thiogalactopyranoside (Biomol Feinchemikalien) was added to a final concentration of 1 mM and incubation was continued for 3 h. The $OD_{550}$ was measured every hour, as was the β-lactamase activity in the culture supernantant to quantify the degree of cell leakiness. Three hours after induction an aliquot of the culture was removed and the cells were lysed by urea lysis followed by lysozyme treatment exactly as described by Knappik and Plückthun (1995). The β-lactamase activity, a marker of the periplasmic fraction, was measured in the culture supernatant, in the insoluble and in the soluble cellular fraction to verify the quality of fractionation and the extent of periplasmic leakiness. Since the leakiness was found to be very low, the results from normalizing to $OD_{550}$ (number of cells) or amount of β-lactamase (accounting for potential loss of periplasmic content to the medium) were identical. The fractions were assayed for antibody fragments by reducing SDS–PAGE, with the samples normalized to $OD_{550}$. The gels were blotted and immunostained using the FLAG antibody M1 (Prickett *et al.*, 1989) as the first antibody, an Fc-specific anti-mouse antiserum conjugated to horseradish peroxidase (Pierce) as second antibody, using a chemiluminescent detection assay described elsewhere (Ge *et al.*, 1995).

### *Purification*

Mutant scFv fragments were purified by a two-column procedure. After French press lysis of the cells, the crude *E.coli* extract was first purified by immobilized metal ion affinity chromatography (IMAC) [Ni-nitrilotriacetic acid (NTA) Superflow, Qiagen] (20 mM HEPES, 500 mM NaCl, pH 6.9; step gradient of imidazole 10, 50 and 200 mM) (Lindner *et al.*, 1992) and, after dialyzing the IMAC eluate against 20 mM MES, pH 6.0, finally purified by cation-exchange chromatography (S-Sepharose fast flow column, Pharmacia) (20 mM MES, pH 6.0; salt gradient 0–500 mM NaCl). Purity was controlled by Coomassie-stained SDS–PAGE. The functionality of the scFv was tested by competition ELISA.

The wild-type (wt) 4-4-20 was expressed as cytoplasmic inclusion bodies in the T7-based system (Studier and Moffatt, 1986; Ge *et al.*, 1995). The refolding procedure was carried out as described elsewhere (Ge *et al.*, 1995). For purification, the refolding solution (2 l) was loaded over 10 h without prior dialysis on to a fluorescein affinity column, followed by a washing step with 20 mM HEPES, 150 mM NaCl, pH 7.5. Two column volumes of 1 mM fluorescein (sodium salt, Sigma), pH 7.5, were used to elute all functional scFv fragments. Extensive dialysis (7 days with 12 buffer changes) was necessary to remove all fluorescein. All purified scFv fragments were tested by gel filtration (Superose-12 column, Pharmacia SMART-System, 20 mM HEPES, 150 mM NaCl, pH 7.5).

### *$K_D$ determination by fluorescence titration*

Protein concentrations were determined photometrically using an extinction coefficient calculated according to Gill and von

Hippel (1989). Fluorescence titration experiments were carried out by taking advantage of the intensive fluorescence of fluorescein. A 2 ml volume of 20 mM HEPES, 150 mM NaCl, pH 7.5, containing 10 or 20 nM fluorescein was placed in a cuvette with an integrated stirrer. The excitation wavelength was 485 nm and emission spectra were recorded from 490 to 530 nm. Purified scFv (in 20 mM HEPES, 150 mM NaCl, pH 7.5) was added in 5–100 µl aliquots and after 3 min of equilibration a spectrum was recorded. All spectra were recorded at 20°C. The maximum of the emission at 510 nM was used for determining the degree of complexation of scFv to fluorescein, seen as quenching, as a function of the concentration of the antibody fragment. The $K_D$ value was determined by Scatchard analysis.

*Equilibrium denaturation measurement*

Equilibrium denaturation curves were obtained by denaturation of 0.2 µM protein in HEPES-buffered saline (HBS) (20 mM HEPES, 150 mM NaCl, 1 mM EDTA, pH 7.5) and increasing amounts of urea (1.0–7.5 M; 20 mM HEPES, 150 mM NaCl, pH 7.4; 0.25 M steps) in a total volume of 1.7 ml. After incubating the samples for 12 h at 10°C and an additional 3 h at 20°C prior to measurements, the fluorescence spectra were recorded at 20°C from 320 to 360 nm with an excitation wavelength of 280 nm. The emission wavelength of the fluorescence peak shifted from 341 to 347 nm during denaturation and was used for determining the fraction of unfolded molecules. Curves were fitted according to Pace (1990).

*Thermal denaturation*

For measuring the thermal denaturation rates, purified scFv was dissolved in 2 ml of HBS buffer to a final concentration of 0.5 µM. The aggregation was followed for 2.5 h at 40°C and at 44°C by light scattering at 400 nm.

*Solubility measurements*

For the solubility determination in HBS buffer (Stevenson and Hageman, 1995), the purified native proteins were concentrated to 1 mg/ml. A stock solution of 40% (w/v) PEG 6000 (Fluka) was prepared in this buffer. A 20 µl volume of protein was incubated for 1 h at room temperature in PEG of various concentrations (5, 8, 10, 15 and 20%) and HBS buffer was added to a final volume of 40 µl. The precipitated scFv fragment was removed by centrifugation at 14 000 r.p.m. (14 500 *g*) in an Eppendorf tube. Coomassie-stained SDS gels of the supernantant were scanned densitometically to estimate the amount of soluble protein as a function of PEG concentration and extrapolated to zero PEG concentration.

## Results

*Comparison of known antibody sequences*

Compared with other domain/domain interfaces in proteins, the interface between immunoglobulin variable and constant domains (v/c interface) is not very tightly packed. A comparison of 30 non-redundant Fab structures in the PDB database showed that between the light chain variable and constant domain an area of 410 ± 90 Å$^2$ per domain is buried, while the heavy chain variable and constant domains interact over an area of 710 ± 180 Å$^2$. Some, but not all, of the v/c interface residues are hydrophobic; predominantly aliphatic. Generally, sequence conservation of the residues contributing to the v/c domain interface is not particularly high. Still, the v/c domain interface shows up as a marked hydrophobic patch on the surface of an Fv fragment (Figure 1).

Solvent-accessible surface areas for 30 non-redundant Fab fragments and their corresponding Fv fragments (derived from the Fab fragment by deleting the constant domain coordinates from the PDB file) were calculated using the program NACCESS (Lee and Richards, 1971). Residues participating in the v/c domain interface were identified by comparing the solvent-accessible surface area of each amino acid side chain in the context of an Fv fragment with its accessible surface in the context of a Fab fragment. Figure 2 shows a plot of the relative change in side-chain accessibility upon deletion of the constant domains as a function of sequence position. Residues which show a significant reduction in side-chain accessibility are also highlighted in the sequence alignment. To assess sequence variability in the positions identified in Figure 2, the variable domain sequences collected in the Kabat database (status March 1996) were analyzed (Table I). Of the 15 interface residues identified in the $V_L$ domain of the antibody 4-4-20 (Figure 1 and Table I), L9(Leu), L12(Pro), L15(Leu), L40(Pro), L83(Leu) and L106(Ile) are hydrophobic and therefore candidates for replacement. Of the 16 interface residues in the $V_H$ domain, H11(Leu), H14(Pro), H41(Pro), H84(Val), H87(Met) and H89(Ile) were identified as possible candidates for substitution by hydrophilic residues in the scFv fragment of the antibody 4-4-20 (Figure 1 and Table I).

Not all of these hydrophobic residues are equally good candidates for replacements, however. While residues which are hydrophobic in one particular sequence but hydrophilic in many other sequences may appear most attractive, the conserved hydrophobic residues listed in Table I have also been investigated, since the evolutionary pressure which kept these conserved residues acted on the Fab fragment within the whole antibody, but not the isolated Fv portion. In this study, we did not consider the proline residues since Pro L40 and Pro H41 form the hairpin turns at the bottom of the framework II region, while the conserved $V_L$ *cis*-proline L8 and proline residues H9 and H14 determine the shape of framework I of the immunoglobulin variable domains.

Excluding prolines, this leaves residues L9 (Leu in 4-4-20, Ser in most κ chains), L15 (Leu, usually hydrophobic), L83 (Leu, usually Val or Phe) and L106 (Ile, as in 86% of all κ chains) in the $V_L$ domain and H11 (Leu, as in 60% of all heavy chains), H84 (Val, in other $V_H$ domains frequently Ala or Ser), H87 (met, usually Ser) and H89 (Ile, most frequently Val) in $V_H$ as possible candidates for replacement in the 4-4-20 scFv fragment.

*Point mutations in the 4-4-20 scFv*

For the particular case of the 4-4-20 scFv fragment, some of the v/c interface residues are already hydrophilic, but nine residues are of hydrophobic nature (including Pro12 in the light chain) (Table I). However, since all permutations of these nine residues with other residues will generate a prohibitive number of experiments, we chose three residues for closer analysis by mutations to test the general concept and biophysical consequences and to define which hydrophobic patch is most important for the aggregation behavior of the wt scFv.

Leu15 in $V_L$ is a hydrophobic amino acid in 98% of all κ chains (Table I). Leu11 is conserved in $V_H$ (Table I) and is involved in v/c interdomain contacts (Lesk and Chothia, 1988). In contrast, valine occurs very infrequently at position H84; mainly found at this position are threonine, serine and alanine (Table I). As can be seen in Figure 1, Val84 contributes to a large hydrophobic patch at the newly exposed surface of $V_H$.

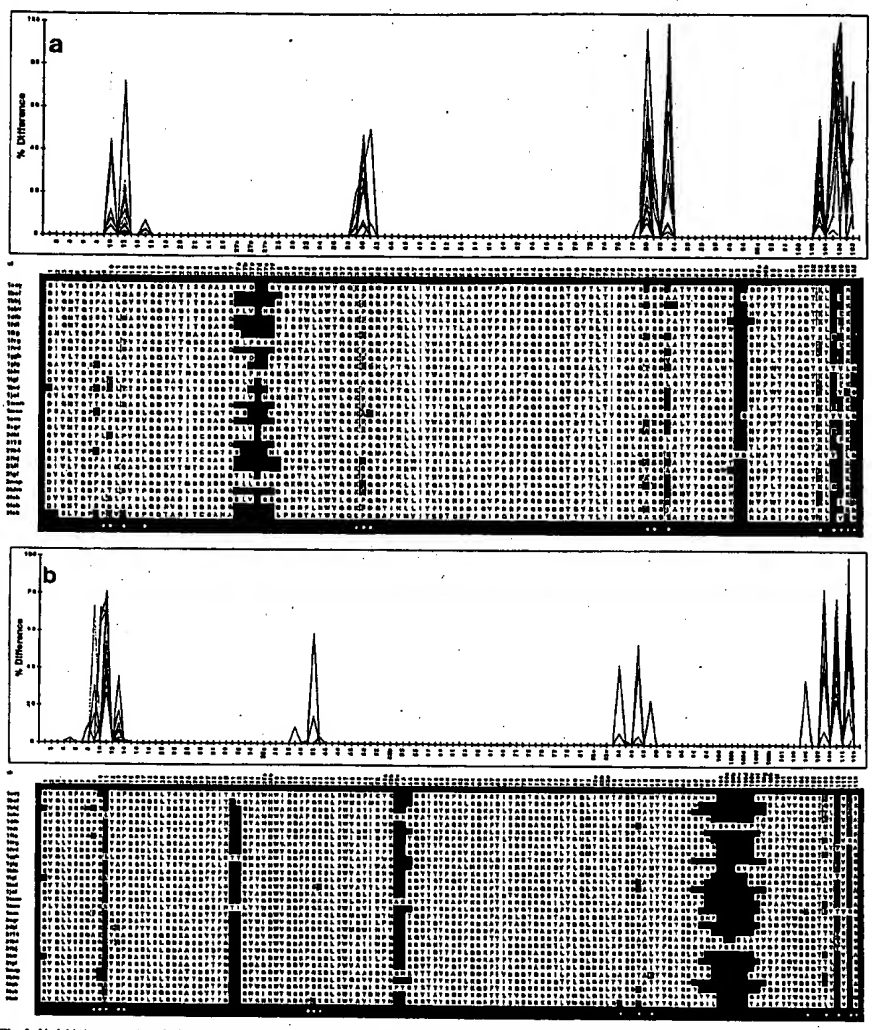
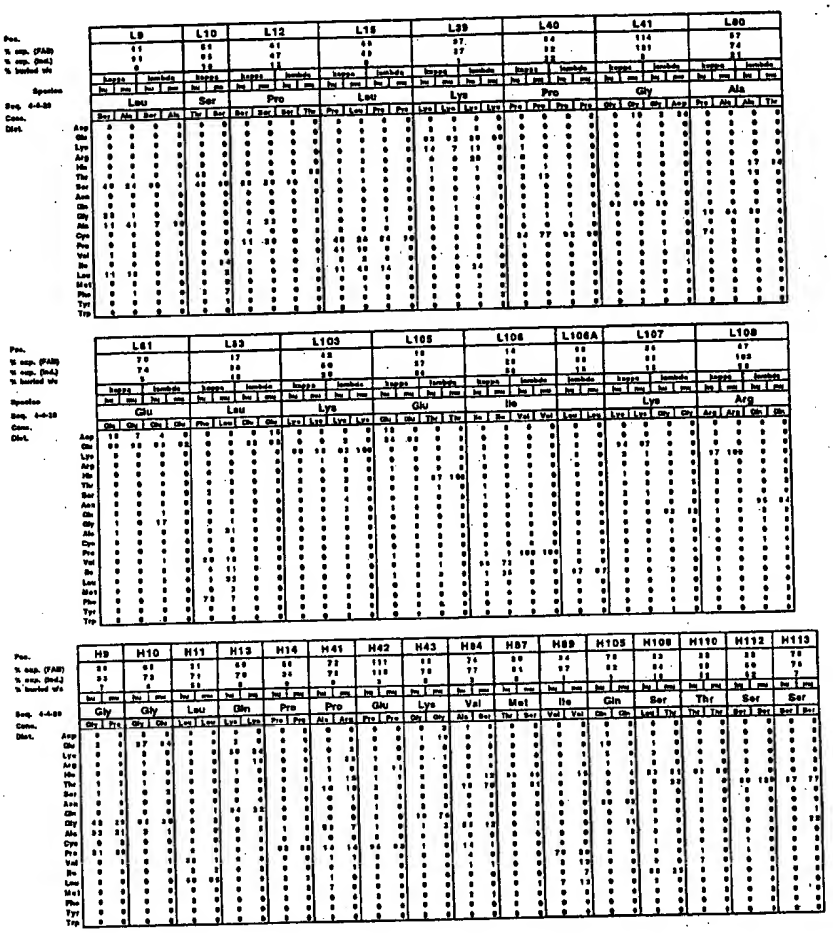



Fig. 2. Variable/constant domain interface residues for (a) $V_L$ and (b) $V_H$. For 30 non-redundant Fab fragments taken from the Brookhaven Databank, the solvent-accessible surface of the amino acid side chains was calculated in the context of an Fv and of an Fab fragment. The plot shows the relative reduction in accessible surface upon contact with the constant domains (color code: 30 different colors for the 30 Fv fragments). In the sequence alignment, residues contributing to the v/c interface are highlighted. The shading from white (<1%) to red (≥80%) reflects the relative reduction of solvent accessible surface upon removing the constant domains (color code: white, <1%; yellow, <20%; yellow–orange, <40%; orange, <60%; red–orange, <80%; red, ≥80%). Circles indicate those positions which are further analyzed in Table I.

Table I. Sequence variability of residues contributing to the v/c interface



Residue statistics are based on the variable domain sequences in the Kabat database (March 1996). Sequences which were <90% complete were excluded from the analysis. Number of sequences analyzed: human VL κ, 404 of 881; murine VL κ, 1061 of 2239; human VL λ, 223 of 409; murine VL λ, 71 of 206; human VH, 663 of 1756; murine VH, 1294 of 3849.

Pos., sequence position according to Kabat *et al.* (1991); % exp (FAB), relative side-chain accessibility (average of 30 non-redundant PDB files) in a Fab fragment as calculated by the program NACCESS [v2.0 by S.Hubbard (http://www.biochem.ucl.ac.uk/~roman/naccess/naccess.html)]; % exp (ind.), relative side-chain accessibility in the isolated VL or VH domain; % buried, relative difference in side-chain accessibility between Fv and Fab fragments; Cons., sequence consensus; Dist., distribution of residue types.

Table II. Mutations introduced in the scFv fragment of the antibody 4-4-20

| | L15E ($V_L$) | L11N ($V_H$) | L11D ($V_H$) | V84D ($V_H$) |
|---|---|---|---|---|
| Flu 1 | ● | | | |
| Flu 2 | | ● | | |
| Flu 3 | | | ● | |
| Flu 4 | | | | ● |
| Flu 5 | | ● | | ● |
| Flu 6 | | | ● | ● |
| Flu 7 | ● | ● | | |
| Flu 8 | ● | | ● | |
| Flu 9 | ● | ● | | |
| Flu 4 short | | | | ● |

Each line represents a different protein carrying the mutations indicated. The residues are numbered according to Kabat *et al.* (1991).

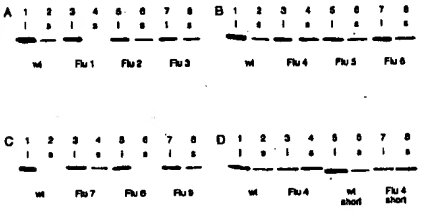



Fig. 3. Western blots showing the insoluble (i) and soluble (s) fractions of cell extracts, prepared as described in Materials and methods, expressing the mutated scFv fragments of the antibody 4-4-20. The amino acids substituted in the various mutants are given in Table II.

All three positions were mutated into acidic residues and L11 was also changed to asparagine (Table II).

The scFv fragment was tested and expressed with two different linkers, the frequently used 15-mer linker $(Gly_4Ser)_3$ (Huston *et al.*, 1995) and the same motif extended to 30 amino acids $(Gly_4Ser)_6$. All mutations were tested in both constructs. The *in vivo* expression results of the different mutations were identical and therefore only the results of the 30-mer linker are described in more detail. The periplasmic expression experiments were carried out at 24.5°C and the ratio of insoluble to soluble (i/s) protein was determined by immunoblotting for every mutant. In Figure 3A–D, lanes 1 and 2, the wt scFv is shown. Almost no soluble protein occurs in periplasmic expression, which is consistent with previous reports by Bedzyk *et al.* (1990) and Denzin *et al.* (1991), who described that the periplasmic expression of wt scFv leads mainly to periplasmic inclusion bodies.

The single-point mutation L15E in $V_L$ (Flu1) shows no effect on the ratio i/s when compared with wt (Figure 3A, lanes 3 and 4). Mutating Leu at position 11 in the heavy chain to asparagine (Flu2) also shows nearly no effect compared with wt, whereas the substitution with aspartic acid (Flu3) changes the i/s ratio to more soluble protein, but still this effect is not very dramatic. In contrast, the point mutation at position 84 (Flu4, Figure 3B, lanes 3 and 4 and Figure 3D, lanes 3 and 4) had a very strong influence on the *in vivo* expression of the scFv fragment of the antibody 4-4-20. The i/s ratio is changed to about 1:1, resulting in a 25-fold increase in soluble expressed protein compared with wt.

The combination of V84D with L11N or L11D (Flu5, Flu6) also changes the i/s ratio compared with wt, but this ratio compared with V84D alone is not improved further (Figure 3B). Interestingly, the combination of Flu5 with the light chain mutation at position 15 (Flu9) leads to less protein in the soluble fraction (Figure 3C, lanes 7 and 8) than Flu5 itself (Figure 3B, lanes and 6). The negative influence of the L15E mutations can also be seen in Flu8 (Figure 3C, lanes 5 and 6) compared with Flu3 (Figure 3A, lanes 7 and 8). In Figure 3D the comparison of the wt (lanes 1, 2 and 5, 6) and Flu4 (lanes 3, 4 and 7, 8) is shown in both the 15-mer and 30-mer constructs. The single-point mutation V84D turned out to be the protein with the best i/s ratio in both constructs, with the 15-mer and the 30-mer linker scFv.

### *Functional expression and purification*

The oligomerization of scFv fragments as a function of linker length has been investigated previously (Holliger *et al.*, 1993; Whitlow *et al.*, 1993, 1994; Desplancq *et al.*, 1994). A continuous decrease in the amount of dimer and multimer formation as a function of linker length has been reported (Desplancq *et al.*, 1994; Whitlow *et al.*, 1994). While the standard $(Gly_4Ser)_3$ linker has been shown to lead to monomeric scFvs in many cases in the $V_H$–$V_L$ direction, this is often not the case in the $V_L$–$V_H$ direction. This is caused by an asymmetry in the $V_L/V_H$ arrangement, leading to a longer distance between the end of $V_H$ and the N-terminus of $V_L$ than between the C-terminus of $V_L$ and N-terminus of $V_H$ (Huston *et al.*, 1995). Consequently, a linker of identical length may lead to different properties of the resulting molecules.

Since we have chosen to use the minimal pertubation FLAG (Knappik and Plückthun, 1994) at the N-terminus of $V_L$ in our constructs and thus the $V_L$–linker–$V_H$ orientation, we have investigated the use of longer linkers. In the periplasmic expression in *E.coli* no difference between the 15-mer and the 30-mer linker in the corresponding mutants is visible (Figure 3D), but when we attempted to purify the two Flu4 scFvs with long and short linker, a discrepancy between the two constructs was found. The purification of the Flu4 mutant (V84D) with the 15-mer linker leads to very small amounts of partially purified protein (~0.015 mg per liter and $OD_{550}$; estimated from SDS–PAGE after IMAC purification), whereas the 30-mer linker construct gives ~0.3 mg per liter and $OD_{550}$ (~1 mg of highly pure functional protein after purification from 1 l normal shake-flask culture). All mutants with 30-mer linker were tested in gel filtration and found to be monomeric (data not shown).

For further *in vitro* characterization three mutants were purified with the 30-mer linker, V84D (Flu4), V84D/L11D (Flu6) and L11D (Flu3). A two-step chromatography, first using IMAC and then cation-exchange chromatography, led to homogeneous protein. The i/s ratio of the antibody fragments (Figure 3) is also reflected in the purification yield of functional protein under these conditions. It should be pointed out that these numbers are only relevant for a comparison of mutants; advances in fermentation technology allow an increase of several hundred-fold in volume yields (Horn *et al.*, 1996). To separate the question of yields per volume from yields per cell, we thus always normalize yields to liter and 1 $OD_{550}$. The mutant Flu4 (V84D) (Figure 3B, lanes 3 and 4) yields ~0.3 mg of purified and functional protein per liter and $OD_{550}$ of cells, Flu6 (L11D/V84D) (Figure 3B, lanes 7 and 8) yields ~0.25 mg per liter and $OD_{550}$ and Flu3 (less protein in the soluble fraction on the blot in Figure 3A, lanes 7 and 8) yields

Table III. $K_D$ values of the different scFv mutants determined in fluorescence titration

| | Flu wt | Flu 3 | Flu 4 | Flu 6 | Flu wt[b] |
|---|---|---|---|---|---|
| $K_D$ (nM)[a] | 26 ± 2.3 | 20 ± 4 | 23 ± 3.3 | 25 ± 4.2 | 90 |

[a]The errors were calculated from the Scatchard analysis.
[b]Whole antibody, determined by Miklasz *et al.* (1995) by intrinsic protein fluorescence.

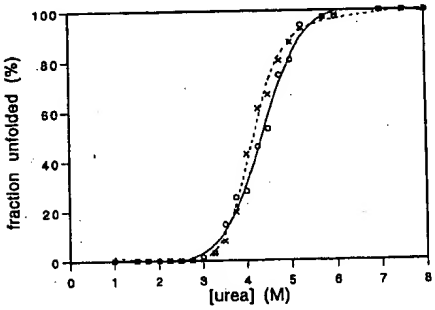



Fig. 4. An overlay plot of the urea denaturation curves is shown. (×) wt scFv; (O) Flu4.

0.05 mg per liter and $OD_{550}$. The wt scFv of the antibody 4-4-20 does not give any functional protein at all in periplasmic expression with either linker and it was therefore expressed as cytoplasmic inclusion bodies, followed by refolding *in vitro* and fluorescein affinity chromatography. The refolded wt scFv was shown by gel filtration to be monomeric with the 30-mer linker (data not shown).

*Biophysical properties of the mutant scFvs*

Since we changed amino acids which are conserved, it could not be excluded *a priori* that changes at these positions may be transmitted through the structure and have an effect on the binding constant, even though they are very far from the binding site (Chatellier *et al.*, 1996). To eliminate this possibility, we determined the binding constant of the mutants Flu3, Flu4 and Flu6 and the wt scFv. Fluorescence titration was used to determine $K_D$ in solution by using the quenching of the intrinsic fluorescence of fluorescein when it binds to the antibody. The fluorescence quenching at 510 nm was measured as a function of added scFv. The $K_D$ values (Table III) obtained for all three mutant scFvs and the wt scFv are very similar and correspond with the recently corrected $K_D$ of the monoclonal antibody 4-4-20 (Miklasz *et al.*, 1995), determined by intrinsic protein fluorescence.

To determine whether the mutations had an influence on the thermodynamic stability of the protein, we determined the equilibrium unfolding curves by urea denaturation. The V84D mutant and the wt scFv were used for this analysis and in Figure 4 an overlay plot is shown. The midpoint of both curves is at 4.1 M urea. Both curves were fitted by an algorithm for a two-state model described by Pace (1990), but the apparent small difference between the V84D mutant and the wt scFv

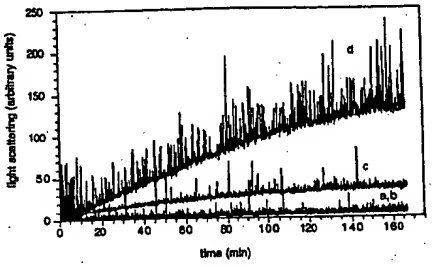



Fig. 5. Thermal denaturation time courses at 40 and 44°C for wt and Flu4 scFv fragments. (a) wt scFv at 40°C; (b) Flu4 at 40°C; (c) Flu4 at 44°C; (d) wt scFv at 44°C.

is not of statistical significance. Furthermore, it is unclear whether a two-state model describes this two-domain protein.

Aggregation of folding intermediates could be an explanation for the different *in vivo* results between the mutant scFvs and the wt scFv (Figure 3). In the periplasm of *E.coli*, the protein concentrations are assumed to be rather high (van Wielink and Duine, 1990) and the aggregation effects could thus be pronounced. In order to estimate the aggregation behavior *in vitro*, we measured the thermal aggregation rates at different temperatures. In Figure 5 it is clearly seen that the wt scFv significantly aggregates already at 44°C, whereas the mutant V84D tends to aggregate more slowly. The wt scFv is thus clearly more aggregation prone than the mutant scFv even though the urea denaturation curves are almost identical. This is very similar to the observations made with different mutations on the antibody McPC603 (Knappik and Plückthun, 1995), where no correlation was found between equilibrium denaturation curves and expression behavior, but a good correlation was found with the thermal aggregation rates.

To determine whether the mutations act on increasing the solubility of the native protein or folding intermediates, we used the solubility in polyethylene glycol (PEG) solutions to determine the solubility of native proteins (Stevenson and Hageman, 1995). PEG can be used to increase the protein concentration above the solubility limit by its excluded volume effect (Middaugh *et al.*, 1979) and is frequently used to induce crystallization. When the concentration of saturated wt and mutant scFv solutions was determined in the presence of various PEG concentrations, after centrifuging precipitated scFv, very similar protein concentrations in the supernantant were obtained for wt and mutant protein (Figure 6). Upon quantification by gel scanning, the solubilities of the two proteins were found to be nearly identical within experimental error and they certainly do not differ by the factor of 25 found in expression yields. Furthermore, both the wt and the Flu4 mutant scFv could be concentrated at least to 10 mg/ml (data not shown). We conclude, therefore, that the mutation must predominantly act on the solubility of folding intermediates and not on that of the native protein.

## Discussion

We have investigated whether there are factors in common to all scFv fragments of antibodies which might make their *in vivo* folding to functional, monomeric proteins inefficient,

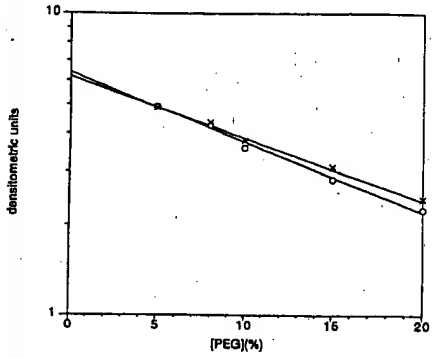



Fig. 6. Solubility of native scFv. (×) wt; (○) Flu4. PEG was used to increase the protein concentration above the solubility limit, precipitated scFv was centrifuged and scFv in the supernatant was analyzed by SDS–PAGE and densitrometric scanning.

with the result of aggregation, but which can be corrected by protein engineering. An analysis of 30 known non-redundant crystal structures showed that there are hydrophobic patches at both former v/c domain interfaces, both in the light chain and in the heavy chain. Since this hydrophobic area has lost its function as a domain interface in an Fv or scFv fragment, the replacement of these hydrophobic surface residues by suitable hydrophilic residues should have no negative influence on folding and stability. Indeed, the disruption of the hydrophobic surface through the substitution of a few key hydrophobic residues by hydrophilic ones can significantly improve the expression and *in vivo* folding behavior of the recombinant Fv and scFv fragments. While the existence of the hydrophobic patches is preserved in all antibodies, their exact position and extent varies. It is thus plausible that the exact nature of these hydrophobic patches may be related to the aggregation properties of these molecules.

We can distinguish conserved residues which are almost always hydrophobic (such as residue 15 in $V_L$ and 11 in $V_H$) and those which can be hydrophilic in a significant number of molecules in the database (such as residue 84 in $V_H$ and 87 in $V_H$). From the present knowledge of aggregation we could not predict *a priori* which of these positions would have a significant influence on the aggregation reaction and should be substituted. Furthermore, the elbow angle (defining the angle between an axis going through $V_H$ and one going through $C_H$, or $V_L$ and $C_L$, respectively) varies widely between different antibody Fab fragments, leading to somewhat different v/c contact surfaces in different antibodies (Wilson and Stanfield, 1994). Therefore, we tested both conserved and less common residues and determined the biophysical properties of the purified scFv to understand the basis of these substitutions. Improvements were seen in both cases, even though the replacement of Val H84 to Asp showed a much stronger effect.

We believe, therefore, that in the optimization of a given antibody fragment, a first inspection of the exposed residues should be carried out according to Table I. Notably those residues which are adjacent to another exposed hydrophobic residue in the structure and allow a common contiguous

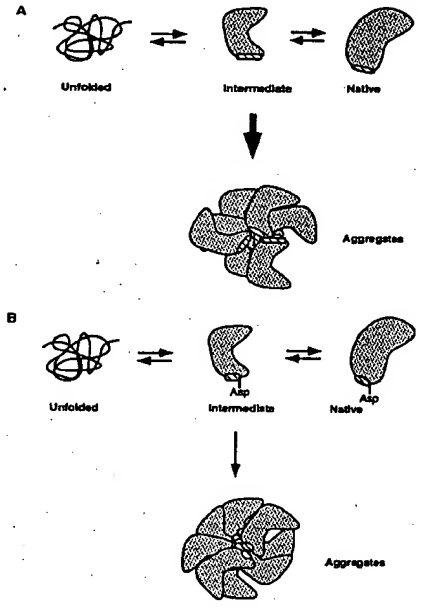



Fig. 7. Schematic folding pathway of the scFvs. (A) wt; (B) Flu4 mutant. The hydrophobic patch is indicating by hatching. There is no evidence of any changes along the productive pathway in rates or equilibrium (horizontally, from left to right). The mutation only seems to inhibit the aggregation reaction. During periplasmic folding, the reaction starts at the top left and proceeds to the top right, with protein being lost to aggregation, according to the thickness of the arrow. Upon thermal denaturation, the reaction starts at the top right and is faster because of a more favored aggregation step in the wt.

interface (such as Val H84 with Met H87 and Ile H89 and to some degree Pro H41; see Figure 1) are candidates for replacement. In contrast, single hydrophobic amino acids such as Leu H11 have a weaker influence, as can be seen in the mutants Flu2 and Flu3.

In the folding pathway of the wt scFv, only a small amount goes to the native form and more than 98% is diverted off the pathway to aggregates (Figure 7A). In the Flu4 scFv the off-pathway reaction is no longer the dominant pathway, when Asp at H84 is introduced (Figure 7B). Since the effect of the mutation is predominantly on folding intermediates and not on the the native protein (Figure 6), the region around H84 may be part of a much bigger hydrophobic surface, possibly involving V/L interface residues or possibly accessible internal ones. The weaker influence of the Leu H11 substitution can be explained by Leu H11 being a single hydrophobic residue on the surface, which may not become part of a contiguous hydrophobic surface in the intermediate and thus does not influence the off-pathway to the same degree. Therefore, also no benefit is visible when both mutations are combined. The difference between Flu2 and Flu3, which is detectable but

small (Figure 3B), may be due to differences in polarity between acidic residues and non-charged residues, as pointed out by Dale *et al.* (1994).

The question then arises to which amino acid types the hydrophobic amino acids should be changed. In general, it may be noted that the isoelectric point of most *E.coli* proteins is acidic (VanBogelen *et al.*, 1992). Thus, the introduction of Asp residues appears particularly attractive, since they also keep the charge close to the main chain. However, it cannot be ruled out that adjacent negative charges are severly destabilizing adjacent loops, as can be seen in the L15E mutation in the light chain (Figures 1 and 3C), where no solubility enhancing effect of the mutation is observed. In this particular case L15 is close to residue E82 in the light chain. Thus, new Asp residues have to be limited to those positions where no such repulsion may occur.

An investigation of the binding constants of the mutants showed that they are identical with the wt scFv within experimental error and thus no measurable conformational change is transmitted through the molecule. While these residues are on the opposite site of the antigen binding site, it is also known that $V_L/V_H$ interface residues, which are about the same distance away from the binding site, can measurably change the binding constant (Chatellier *et al.*, 1996), probably by domain reorientation. The residues targeted here do not seem to have any long-range effects and are not expected to change the domain orientation.

The mutations described here do fall into the same class as those described earlier for an scFv fragment of the antibody McPC603 (Knappik and Plückthun, 1995), green fluorescent protein (Crameri *et al.*, 1996) and glutathione reductase (Leistler and Perham, 1994) in that they do not affect the free energy of folding, as measured in urea denaturation experiments, but lower the aggregation rate *in vitro*. Thus, the large amount of insoluble wt protein *in vivo* does not result from an unfolding of an unstable native protein, because the mutants are not more stable (Figure 4), but give more soluble protein. We can also exclude the aggregation of native protein *in vivo* as the cause (Figure 6), since the solubility of the native mutant proteins is identical. Thus we propose that this aggregation occurs as an off-pathway in the folding reaction. The decisive factor is the aggregation rate itself, which differs between the mutants (Knappik and Plückthun, 1995). These mutants are thus of a different nature from the mutants of REI (Chan *et al.*, 1996), where inclusion body formation and $\Delta G$ are inversely correlated.

Currently, there are three strategies available to improve the folding behavior of antibodies by protein engineering: (i) loop grafting to superior frameworks, (ii) back engineering from sequence comparisons and (iii) interface engineering. Clearly, these approaches are not mutually exclusive and their various implications will be briefly discussed.

The humanization of the anti-HER2 antibody 4D5 and thus the grafting to this human framework was found to be responsible for the superior expression properties of this Fab fragment, since the corresponding murine Fab fragment, in the identical *E.coli* host–vector system, gave rise to a 100-fold lower expression yield under identical fermentation conditions (Carter *et al.*, 1992). Loop grafting to superior frameworks must overcome several challenges, however, which again are only starting to be understood. It must be elucidated whether a framework is particularly robust against aggregation and how independent this is of the CDRs. Moreover, numerous contacts between CDRs and framework are important in mediating the exact conformation of all CDRs (Chothia *et al.*, 1989; Tramontano *et al.*, 1990), restricting the choice of some amino acids.

Back engineering from sequence comparisons (Knappik and Plückthun, 1995) has been a valuable tool in establishing that indeed single residues can decide the fate of the protein and was originally inspired from the loop-grafting experiments mentioned above. However, many more experiments will be necessary before predictive rules will emerge. Since every given antibody may have different problems, these solutions will be tailor-made for a given antibody. The more important point is, however, that these rules or sequences emerging can directly be used for improving the general frameworks to be used in synthetic libraries.

While loop grafting and sequence back engineering can be carried out in all common antibody formats (Fab, Fv, scFv) and have been shown to have beneficial effects in all formats, the interface engineering will depend on the desired antibody format. The Fab fragment is a complete unit within the antibody, separated from the rest of the molecule by a flexible hinge region with no further protein–protein contacts, but any subdivision of the Fab fragment will expose a new interface. When using $V_H$ domains by themselves (Ward *et al.*, 1989), the large $V_H/V_L$ interface is exposed, frequently resulting in insoluble molecules (Kortt *et al.*, 1995). Camels (Desmyter *et al.*, 1996) and Llamas (Spinelli *et al.*, 1996) possess a subset of antibodies with unpaired soluble $V_H$ domains and the crystal structures show that the residues normally forming the $V_H/V_L$ interface are more hydrophilic. Indeed, it is possible to introduce these mutations in a human $V_H$ domain (Davies and Riechmann, 1994, 1996) to increase solubility. However, while the domains become less 'sticky', they have lower melting points than the corresponding original human domains (Davies and Riechmann, 1996) even after adding stabilizing mutations. The $V_H/V_L$ interface mutations are distinct from the v/c mutations investigated in the present study, which also do not cause a loss in stability.

The engineering of the v/c interface is in principle a general tool which can be applied both to general frameworks for libraries or individual antibody sequences. It was conceived from the notion that all Fv and scFv fragments have in common an unnatural former interface. A similar analysis has been carried out for scTCRs (Novotny *et al.*, 1991); however, different residues were mutated. We have identified all candidate residues at the former v/c domain interface and we have shown that this mutagenesis strategy can have a very strong effect. It will be very useful to avoid these hydrophobic patches in a synthetic scFv framework library, since they no longer serve any purpose. In such an optimized framework, adventitious exposed hydrophobic residues, which may arise in a random mutagenesis approach (Palzkill and Botstein, 1992; Stemmer, 1994) would be much less harmful since they would not likely be part of a contiguous hydrophobic surface. Thus, a short-term solution would be to use Table I and Figure 1 as a guide to spot rapidly hydrophobic patches in given sequences.

In the course of these studies we also reinvestigated the use of the standard $(Gly_4Ser)_3$ linker. A number of reports have shown that scFv fragments dimerize or multimerize with shorter linkers (Holliger *et al.*, 1993; Whitlow *et al.*, 1993, 1994; Desplancq *et al.*, 1994; Wu *et al.*, 1996). We compared a 15- and a 30-mer linker and found that the relative influence

of the mutations described is independent of the linker length. The difference in purification is due to oligomerization of the scFv fragment with the 15-mer linker (data not shown). The 30-mer linker we used seems to avoid diabody formation or multimerization. In general, we found 20-mer linkers to be sufficient to yield monomers reliably (Plückthun *et al.*, 1996).

In conclusion, we have described a very general problem of scFv fragments in that they expose a former interface with hydrophobic patches and that strategic placement of negative charges can very significantly improve the *in vivo* folding properties of these molecules. Most scFv fragment used in biomedical research have to be produced in large amounts and thus the described strategy and the use of Figure 1 and Table I may serve as a guide to improve the expression level. If monoclonal antibodies will be replaced by synthetic libraries, such findings should be incorporated into the design of general frameworks, since the described hydrophobic residues no longer serve any purpose in the scFv fragments. Finally, this strategy may also be useful for other engineered protein domains, which are taken out of their natural context, to improve their folding properties and the *in vivo* expression yield.

## Acknowledgements

We thank Dr Karl Proba for helpful discussions. This work was financially supported by the Schweizerische Nationalfonds, grant 3100-037717.93/1.

## References

Bedzyk,W.D., Weidner,K.M., Denzin,L.K., Johnson,L.S., Hardman,K.D., Pantoliano,M.W., Asel,E.D. and Voss,E.W.,Jr (1990) *J. Biol. Chem.*, 265, 18615–18620.

Bird,R.E., Hardman,K.D., Jacobson,J.W., Johnson,S., Kaufman,B.M., Lee,S.M., Lee,T., Pope,S.H., Riordan,G.S. and Whitlow,M. (1988) *Science*, 242, 423–426.

Brinkmann,U., Reiter,Y., Jung,S., Lee,B. and Pastan,I. (1993) *Proc. Natl Acad. Sci. USA*, 90, 7538–7542.

Carter,P. *et al.* (1992) *Bio/Technology*, 10, 163–167.

Chan,W., Helms,L.R., Brooks,I., Lee,G., Ngola,S., McNulty,D., Maleeff,B., Hensley,P. and Wetzel,R. (1996) *Folding Des.*, 1, 77–89.

Chatellier,J., van Regenmortel,M.H.V., Vernet,T. and Altschuh,D. (1996) *J. Mol. Biol.*, 264, 1–6.

Chothia,C. *et al.* (1989) *Nature*, 342, 877–883.

Crameri,A., Whitehorn,E.A., Tate,E. and Stemmer,W.P.C. (1996) *Nature Biotechnol.*, 14, 315–319.

Dale,G.E., Broger,C., Langen,H., D'Arcy,A. and Stüber,D. (1994) *Protein Engng*, 7, 933–939.

Davies,J. and Riechmann,L. (1994) *FEBS Lett.*, 339, 285–290.

Davies,J. and Riechmann,L. (1996) *Protein Engng*, 9, 531–537.

Denzin,L.K., Whitlow,M. and Voss,E.W.,Jr (1991) *J. Biol. Chem.*, 266, 14095–14102.

Desmyter,A., Transue,T.R., Ghahroudi,M.A., Thi,M.D., Poortmans,F., Hamers,R., Muyldermans,S. and Wyns,L. (1996) *Nature Struct. Biol.*, 3, 803–811.

Desplancq,D., King,D.J., Lawson,A.D. and Mountain,A. (1994) *Protein Engng*, 7, 1027–1033.

Ge,L., Knappik,A., Pack,P., Freund,C. and Plückthun,A. (1995) In Borrebaeck,C.A.K. (ed.), *Antibody Engineering. A Practical Approach*. IRL Press, Oxford, pp. 229–266.

Hill,S.C. and von Hippel,P.H. (1989) *Anal. Biochem.*, 182, 319–326.

Hlockshuber,R., Malia,M., Pfitzinger,I. and Plückthun,A. (1990) *Biochemistry*, 29, 1362–1366.

Holliger,P., Prospero,T. and Winter,G. (1993) *Proc. Natl Acad. Sci. USA*, 90, 6444–6448.

Horn,U. *et al.* (1996) *Appl. Microbiol. Biotechnol.*, 46, 524–532.

Huston,J.S., Levinson,D., Mudgett-Hunter,M., Tai,M.S., Novotny,J., Margolies,M.N., Ridge,R.J., Bruccoleri,R.E., Haber,E., Crea,R. and Oppermann,H. (1988) *Proc. Natl Acad. Sci. USA*, 85, 5879–5883.

Huston,J., George,A.J.T., Tai,M., McCartney,J.E., Jin,D., Segal,D.M., Keck,P. and Oppermann,H. (1995) In Borrebaeck,C.A.K. (ed.), *Antibody Engineering. A Practical Approach*. IRL Press, Oxford, pp. 185–228.

Kabat,E.A., Wu,T.T., Perry,H.M., Gottesmann,K.S. and Foeller,C. (1991) *NIH Publication 91-3242*. US Department of Health and Human Services, Public Health Service, National Institutes of Health, Washington, DC.

Knappik,A. and Plückthun,A. (1994) *BioTechniques*, 17, 754–761.

Knappik,A. and Plückthun,A. (1995) *Protein Engng*, 8, 81–89.

Korti,A.A., Guthrie,R.E., Hinds,M.G., Power,B.E., Ivancic,N., Caldwell,J.B., Gruen,L.C., Norton,R.S. and Hudson,P.J. (1995) *J. Protein Chem.*, 14, 167–178.

Kunkel,T.A., Roberts,J.D. and Zakour,R.A. (1987) *Methods Enzymol.*, 154, 367–382.

Lee,B. and Richards,F.M. (1971) *J. Mol. Biol.*, 55, 379–400.

Lesk,A.M. and Chothia,C. (1988) *Nature*, 335,188–190.

Leistler,B. and Perham,R.N. (1994) *Biochemistry*, 33, 2773–2781.

Lindner,P., Guth,B., Wülfing,C., Krebber,C., Steipe,B., Müller,F. and Plückthun,A. (1992) *Methods: Companion Methods Enzymol.*, 4, 41–56.

Middaugh,C.R., Tisel,W.A., Haire,R.N. and Rosenberg,A. (1979) *J. Biol. Chem.*, 254, 367–370.

Miklasz,S.D., Gulliver,G.A. and Voss,E.W.,Jr (1995) *J. Mol. Recogn.*, 8, 258–269.

Novotny,J., Ganju,R.K., Smiley,S.T., Hussey,R.E., Luther,M.A., Reeny,M.A., Siliciano,R.F. and Reinherz,E.L. (1991) *Proc. Natl Acad. Sci. USA*, 88, 8646–8650.

Pace,C.N. (1990) *Trends Biotechnol.*, 8, 93–98.

Pajzkill,T. and Botstein,D. (1992) *Proteins*, 14, 29–44.

Plückthun,A. (1992) *Immunol. Rev.*, 130,151–188.

Plückthun,A. (1994) In van Oss,C.J. and van Regenmortel,M.H.V. (eds), *Immunochemistry*. Marcel Dekker, New York, pp. 201–235.

Plückthun,A., Krebber,A., Krebber,C., Horn,U., Knüpfer,U., Wenderoth,R., Nieba,L., Proba,K. and Riesenberg,D. (1996) In McCafferty,J. (ed.), *Antibody Engineering. A Practical Approach*. IRL Press, Oxford, pp. 203–252.

Prickett,K.S., Amberg,D.C. and Hopp,T.P. (1989) *BioTechniques*, 7, 580–589.

Prodromou,C. and Pearl,L.H. (1992) *Protein Engng*, 5, 827–829.

Spinelli,S., Frenken,L., Bourgeois,D., Ron,L., Bos,W., Verrips,T., Anguille,C., Cambillau,C. and Tegoni,M. (1996) *Nature Struct. Biol.*, 3, 752–757.

Stemmer,W.P.C. (1994) *Nature*, 370, 389–391.

Stevenson,C.L. and Hageman,M.J. (1995) *Pharm. Res.*, 12, 1671–1676.

Studier,F.W. and Moffatt,B.A. (1986) *J. Mol. Biol.*, 189, 113–130.

Tramontano,A., Chothia,C. and Lesk,A. M. (1990) *J. Mol. Biol.*, 215, 175–182.

Ullrich,H.D., Patten,P.A., Yang,P.L., Romesberg,F.E. and Schultz,P.G. (1995) *Proc. Natl Acad. Sci. USA*, 92, 11907–11911.

VanBogelen,R.A., Sankar,P., Clark,R.L., Bogan,J.A. and Neidhardt,F.C. (1992) *Electrophoresis*, 13, 1014–1054.

Van Wielink,J.E. and Duine,J.A. (1990) *Trends Biochem. Sci.*, 15, 136–137.

Ward,E.S., Güssow,D., Griffiths,A.D., Jones,P.T. and Winter,G. (1989) *Nature*, 341, 544–546.

Whitlow,M. *et al.* (1993) *Protein Engng*, 6, 989–995.

Whitlow,M., Filupa,D., Rollence,M.L., Feng,S. and Wood,J.F. (1994) *Protein Engng*, 7, 1017–1026.

Whitlow,M., Howard,A.J., Wood,J.F., Voss,E.W.,Jr and Hardman,K.D. (1995) *Protein Engng*, 8, 749–761.

Wilson,I.A. and Stanfield,R.L. (1994) *Curr. Opin. Struct. Biol.*, 4, 857–867.

Winter,G. and Milstein,C. (1991) *Nature*, 349, 293–299.

Wu,A.M. *et al.* (1996) *Immunotechnology*, 2, 21–36.

Young,N.M., MacKenzie,C.R., Narang,S.A., Oomen,R.P. and Baenziger,J.E. (1995) *FEBS Lett.*, 377, 135–139.

*Received September 12, 1996; revised November 8, 1996; accepted November 19, 1996*